

PACIFIC LINGUISTICS

Series C - No. 94

FOCAL II: PAPERS FROM THE
FOURTH INTERNATIONAL CONFERENCE ON
AUSTRONESIAN LINGUISTICS

*

Paul Geraghty
Lois Carrington
S.A. Wurm
eds



Department of Linguistics
Research School of Pacific Studies
THE AUSTRALIAN NATIONAL UNIVERSITY

PACIFIC LINGUISTICS is issued through the Linguistic
Circle of Canberra and consists of four series:

SERIES A - Occasional Papers
SERIES B - Monographs
SERIES C - Books
SERIES D - Special Publications

EDITOR: S.A. Wurm

ASSOCIATE EDITORS: D.C. Laycock, C.L. Voorhoeve, D.T. Tryon, T.E. Dutton

EDITORIAL ADVISERS:

B.W. Bender University of Hawaii	K.A. McElhanon Summer Institute of Linguistics
David Bradley La Trobe University	H.P. McKaughan University of Hawaii
A. Capell University of Sydney	P. Mühlhäusler Linacre College, Oxford
Michael G. Clyne Monash University	G.N. O'Grady University of Victoria, B.C.
S.H. Elbert University of Hawaii	A.K. Pawley University of Auckland
K.J. Franklin Summer Institute of Linguistics	K.L. Pike Summer Institute of Linguistics
W.W. Glover Summer Institute of Linguistics	E.C. Polomé University of Texas
G.W. Grace University of Hawaii	Malcolm Ross Australian National University
M.A.K. Halliday University of Sydney	Gillian Sankoff University of Pennsylvania
E. Haugen Harvard University	W.A.L. Stokhof University of Leiden
A. Healey Summer Institute of Linguistics	B.K. T'sou City Polytechnic of Hong Kong
L.A. Hercus Australian National University	E.M. Uhlenbeck University of Leiden
Nguyễn Đăng Liêm University of Hawaii	J.W.M. Verhaar Divine Word Institute, Alexishafen
John Lynch University of Papua New Guinea	

All correspondence concerning PACIFIC LINGUISTICS, including
orders and subscriptions, should be addressed to:

The Secretary
PACIFIC LINGUISTICS
Department of Linguistics
Research School of Pacific Studies
The Australian National University
Canberra, A.C.T. 2601
Australia.

Copyright © The Authors

First Published 1986

Typeset by Sue Tys

Printed by A.N.U. Printing Service

Bound by Adriatic Bookbinders Pty Ltd

Maps drawn by Theo Baumann, Cartography, Department of Linguistics, Research
School of Pacific Studies, Australian National University.

The editors are indebted to the Australian National University for assistance in the
production of this series.

This publication was made possible by an initial grant from the Hunter Douglas Fund.

National Library of Australia Card Number and ISBN for complete set of two
volumes: 0 85883 343 3

National Library of Australia Card Number and ISBN for this volume: 0 85883 345 X

TABLE OF CONTENTS

<i>FOCAL II: THE CONTRIBUTORS</i>		iv
Further thoughts on Oceanic subgrouping	George W. GRACE	1-12
On the nature of subgrouping arguments	S.P. HARRISON	13-21
Proto-Austronesian laterals and nasals	Isidore DYEN and Shigeru TSUCHIDA	23-31
Linguistic variations of different age groups in some Formosan languages	Paul Jen-kuei LI	33-50
Lexical diffusion in Sangir	J.N. SNEDDON	51-71
The Samosir dialect of Toba-Batak	J.P. SARUMPAET	73-86
The Barrier island languages in the Austronesian language family	Bernd NOTHOFER	87-109
Some problems of diachronic typology of the Malayo-Javanic languages	A.K. OGLOBLIN	111-122
Eastern Seram: a subgrouping argument	James T. COLLINS	123-146
The genetic relationships of Philippine languages	R. David ZORC	147-173
A genetic grouping of Oceanic languages in Bougainville and the Western Solomons	Malcolm ROSS	175-200
On determining the external relationships of the Micronesian languages	Frederick H. JACKSON	201-238
Stem-initial consonant alternation in the languages of Epi, Vanuatu: a case of assimilation?	D.T. TRYON	239-258
The Proto-Southern Vanuatu pronominal system	John LYNCH	259-287
The sound system of Proto-Central-Pacific	Paul GERAGHTY	289-312
The chronology of three Samoan sound changes	Even HOVDHAUGEN	313-331
Linguistic convergence in Central Vanuatu	Ross CLARK	333-342
Counting and culture contact in north-east New Guinea	Geoff P. SMITH	343-349
Police Motu and the Second World War	Tom DUTTON	351-406
Refugees, traders, and other wanderers: the linguistic effects of population mixing in Melanesia	Ann CHOWNING	407-434
Wai: a Malaitan language in Fiji	Jeff SIEGEL	435-463

FOCAL II: THE CONTRIBUTORS

ANN CHOWNING was born in Arkansas and educated in Pennsylvania (BA in Spanish, Bryn Mawr College, MA and PhD in Anthropology, University of Pennsylvania). Since 1954 she has been carrying out long-term anthropological fieldwork in Papua New Guinea, working in four Austronesian-speaking societies, three in West New Britain and one in Milne Bay Province. She is particularly interested in linguistic evidence for the culture history of the Pacific. Ann Chowning is Professor of Anthropology at Victoria University, Wellington, New Zealand.

ROSS CLARK was born in Vancouver, British Columbia. He graduated from the University of British Columbia, and then studied linguistics at Cornell University and the University of California, San Diego, from which he received his PhD in 1973. He is Senior Lecturer in Linguistics in the Department of Anthropology, University of Auckland. His major research interests are in the languages of Central Vanuatu, Oceanic comparative and historical linguistics, and language contact.

JAMES T. COLLINS is interested especially in the languages of Central Maluku. In 1980 he received his PhD in linguistics from the University of Chicago and began lecturing in historical linguistics and Malay dialectology at the National University of Malaysia. Since 1984 he has been Professor of Indonesian at the University of Hawaii. He has worked on the revision of the Echols-Shadily Indonesian dictionary, and is consultant to Malaysia's national dialectology project as well as the History of Malay project.

TOM DUTTON is a Senior Fellow in the Department of Linguistics, Research School of Pacific Studies, at the Australian National University, where he received his PhD (for a study of Koiarian languages) in 1969. His main interests are in the two major lingue franche of Papua New Guinea, Tok Pisin and Police Motu, and in comparative historical descriptive studies in non-Austronesian languages in south-east mainland Papua.

ISIDORE DYEN was until June 1984 Professor of Comparative Linguistics at Yale University; he is now Adjunct Professor in Linguistics at the University of Hawaii, Honolulu. He maintains his interest in Austronesian comparative linguistics and reconstruction.

PAUL GERAGHTY was born in England of Irish parents, and educated at Rugby School and Cambridge University. His doctoral dissertation, from the University of Hawaii, was on the history of the Fijian languages. He has been working since 1978 as Consultant/Researcher at the Fijian Dictionary Project, in Suva.

GEORGE W. GRACE is Professor of Linguistics at the University of Hawaii. His research interests include the history of Austronesian (especially Oceanic) linguistics, linguistic diachrony, and the nature of language generally.

SHELDON HARRISON is on the faculty of the Department of Anthropology of the University of Western Australia. His research interests include the languages of Kiribati, and Mokilese and other Micronesian languages.

EVEN HOVDHAUGEN has been Professor of General Linguistics at the University of Oslo since 1974. He has published a number of books and articles in the fields of general linguistics, comparative Indo-European philology, Turcology and the history of linguistics.

FREDERICK H. JACKSON received his PhD in Linguistics, and MA in English as a Second Language, at the University of Hawaii at Manoa, where he worked for several years as staff member and then Coordinator of the Bilingual Education Program for Micronesia. Since leaving Hawaii, he has taught at universities in Thailand and Pennsylvania, and is now a Language Training Supervisor at the School of Language Studies of the US Foreign Service in Washington DC. His primary interests include comparative Oceanic and Tai linguistics, theoretical phonology, and second-language acquisition and teaching.

PAUL JEN-KUEI LI comes from Taiwan. He received an MA in English at the University of Michigan in 1963, and a PhD in Linguistics at the University of Hawaii in 1973. He taught at the National Taiwan Normal University 1963-1967, Nanyang University 1973-1974, National Taiwan University 1974-1984, the University of Hawaii (Fall 1984) and the Ohio State University (Summer 1985). He is currently Research Fellow and Chairman of the Linguistics Section at the Institute of History and Philology, Academia Sinica in Taipei, Taiwan. His main interests are in linguistic theories and historical reconstruction.

JOHN LYNCH is originally from Sydney, Australia. After taking an undergraduate degree in Anthropology at the University of Sydney, he studied at the University of Hawaii, where he received his PhD in linguistics in 1974. He has been on the staff of the University of Papua New Guinea since 1978, and is currently Professor of Language and Literature. His main interests are in comparative Austronesian linguistics, Melanesian culture history, pidgins and creoles, and language policy.

BERND NOTHOFER comes from Usingen, West Germany; he completed a doctorate in Austronesian linguistics at Yale University. He is currently Professor of Southeast Asian Studies at the Johann Wolfgang Goethe-Universität, Frankfurt; his main interests are in comparative Austronesian and descriptive Indonesian linguistics.

ALEKSANDR K. OGLOBLIN graduated from the Oriental Faculty of Leningrad University. After several years in the army as interpreter he has been teaching Indonesian language and literature in the same Faculty, where he is now a Senior Lecturer. His PhD dissertation was on causative verbs in Indonesian. He is active in the field of grammar and phonology of Malay, Madurese and Javanese. He is writing a monograph on the structure of Malayo-Javanic languages and their typological evolution.

MALCOLM D. ROSS, formerly Principal of the Goroka campus of the University of Papua New Guinea, has completed a PhD thesis on the genetic relationships of the Oceanic languages of Papua New Guinea, at the Australian National University, where he is now engaged in research into comparative Austronesian linguistics.

J.P. SARUMPAET is Senior Lecturer in Indonesian in the Department of Indian and Indonesian Studies, at the University of Melbourne. His special interest is in the dialects of Toba-Batak; he is working on a dictionary of Toba-Batak.

JEFF SIEGEL was born in Chicago but has spent most of the last 15 years in the Pacific, mainly teaching in Fiji, New Zealand, and Papua New Guinea. He completed his PhD in linguistics at the Australian National University, and is now a fellow in language studies at the Pacific Language Unit of the University of the South Pacific in Port Vila, Vanuatu.

GEOFF P. SMITH is from Northumberland, UK. He studied at the Universities of Bristol, Sussex and Papua New Guinea, and has taught in England, Ghana and Papua New Guinea. He is now a Senior Lecturer in the Department of Language and Social Science at the University of Technology, Lae, Papua New Guinea. Interests include languages and counting systems of Morobe Province, reading strategies for learners of English as a second language, and creolised Tok Pisin.

JAMES N. SNEDDON graduated from the University of Sydney and completed his doctoral thesis "Tondano phonology and grammar" at the Australian National University. He is at present Lecturer in Indonesian Language at Griffith University, Brisbane. His main research interest is in comparative Indonesian linguistics.

DARRELL TRYON is a Senior Fellow in the Department of Linguistics, Research School of Pacific Studies, Australian National University. He has worked on the description and classification of languages of Vanuatu, Loyalty Islands, Tahiti, Solomon Islands, and also Pacific pidgins, and is engaged in a general survey of the languages of the Pacific.

SHIGERU TSUCHIDA is Professor of Linguistics in the Department of Linguistics of the University of Tokyo. His chief research interests are concerned with the Austronesian languages of Taiwan and the Philippines, on which he has published widely.

R. DAVID ZORC comes from North Chicago, Illinois; after work with the Peace Corps (1965-1969) in the Philippines, he undertook his doctorate at Cornell University (1969-1973) and worked as an Assistant in Research at Yale University (1973-1975). Since 1976 he has been a Senior Lecturer in the School of Australian Linguistics of the Darwin Institute of Technology (formerly Darwin Community College), Northern Territory of Australia. His main interests are lexicography, subgrouping, and historical reconstruction.

FURTHER THOUGHTS ON OCEANIC SUBGROUPING

George W. Grace

My involvement with Oceanic languages, which began 30 years ago, was associated from the outset with the problem of subgrouping.¹ It came about because some anthropologists who were interested in reconstructing the culture history of the Pacific area wanted an account of the genetic relationships within the Austronesian family and particularly within the Oceanic branch of the family. My remarks here will be limited to the Oceanic languages although the approach which I follow should be applicable to the rest of Austronesian, or for that matter, to any other language family.

I have avoided saying that what these anthropologists wanted was a 'family tree' or a 'subgrouping', although I am sure that they would have found either characterisation acceptable. I have chosen instead the more general statement 'an account of the genetic relationships' to indicate that I think their needs could have been satisfied with some other kind of account.

In fact, work going back as far as Isidore Dyen's lexicostatistical classification of the Austronesian languages (Dyen 1965) has recognised patterns of relationship different from those recognised in the classical formulation of the family tree model. Recognition of such relationships has appeared regularly in recent discussions by Paul Geraghty, Andrew Pawley,² and most of the other scholars who have been involved in research on genetic relationships among Oceanic languages. However, most of these discussions have continued to use family tree *terminology*. That is, such terms as "proto-language" and "subgroup" have been used, but in loosened senses. What I want to suggest here is that the real lesson in all of this recent work is precisely that the family tree model is not adequate to represent the genetic relationships among the Oceanic languages. And I suggest furthermore that by continuing to use the family tree terminology we inhibit ourselves from realising the full benefits of this lesson.

What we are concerned with are the cases where a given member of the Oceanic group is more closely related to some of its co-members than to others. The established practice has been to conceive of these closer relationships in terms of subgroups in a family tree. But, as I will try to show here, it is possible to separate the concept of closer relationship from the concept of subgrouping and the assumptions of the family tree model.³ Accordingly, it will be possible to claim that two members, A and B, of a language family are in a closer relationship with each other than with some other member, C, without claiming that there

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 1-12. *Pacific Linguistics*, C-94, 1986.

© George W. Grace

is a well-defined subgroup of the family to which those two belong. To say that A and B are in a closer relationship as opposed to C will mean nothing more than that they have undergone one or more innovations in common which C has not undergone.

Let me begin by recalling that the family tree model is not the only recognised model of linguistic differentiation. In fact, it is a standard practice in textbook treatments of linguistic differentiation to present the family tree (Stammbaum) model of August Schleicher and the wave (Wellentheorie) model of Johannes Schmidt as two competing models, neither of which has succeeded in displacing the other.

Now there is in this what seems to me to be a strange contradiction. It seems to be accepted as a truism that the wave model corresponds to the actual process of differentiation more accurately than does the family tree model. On the other hand, when someone wants an account of the genetic relationships within a language family, the automatic expectation seems to be that the answer is to be expressed in terms of the family tree model. It is appropriate to ask why that should be so.

In what follows I will briefly consider some factors which may help explain the popularity of the family tree model. Then I will try to show that the wave model is indeed a more accurate representation of the kinds of events which must have led from the Oceanic proto-language to the contemporary linguistic situation. Finally, I will attempt a very rough account, based on the wave model, of the probable historical relationships among the Oceanic languages.

ATTRactions OF THE FAMILY TREE MODEL

The question is: If the family tree model so seriously fails to reflect the actual historical relationships, why has it persisted in use to the extent that it has? I cannot attempt to give a complete answer, but I can make a few suggestions.

I believe that there are probably two considerations which have tended to encourage its use. One consideration which has surely been significant is that it has a visually attractive product. A family tree offers the great convenience of being two-dimensional, it can be printed on a single sheet of paper, and it can be easily read. This is certainly a very attractive feature of the approach. However, I would suggest that this aesthetic appeal is bought at the cost of accuracy.

A second consideration, I suspect, is that this approach to linguistic relationships heightens the analogy between genetic linguistics and evolutionary biology. The point of analogy which has always seemed to me to be particularly striking is the parallel role attributed to particular key evolutionary units in each. The units of which I am speaking are, respectively, the species in biology (as contrasted with subspecies or races) and the individual language in linguistics (as contrasted with dialects, etc.).

What is of such crucial importance about these units is that they are supposed to represent the point-of-no-return in the process of differentiation. Evolutionary biology (or, respectively, historical linguistics) traditionally makes the following assumptions. Subspecies (or dialects) may arise through natural differentiation, but unless they are physically isolated from the other subspecies (or dialects), gene flow (or the spread of innovations) will continue

among them. As long as this gene flow (or spread of innovations) continues, a levelling of the differences remains possible.

However, there is assumed to be a sort of biological (or linguistic) Rubicon which is occasionally crossed. This Rubicon-crossing is supposed to occur when the subspecies boundary is elevated to the status of a species boundary (or when a dialect boundary is elevated to the status of a language boundary). From this point on, the newly established unit - the species or the language - has become irreversibly separate from its relatives. Gene flow (or the spread of linguistic innovations) is no longer possible.⁴

My objection to this biological analogy is that, as far as linguistics is concerned, it is a false picture. I do not believe that there is any such Rubicon-crossing between dialect differences on the one hand and language differences on the other. I would propose a different picture of the process of differentiation.

I would say that any local speech variety will produce linguistic innovations. Some innovations will spread from one variety to its neighbours. Two conditions are necessary if an innovation is to spread from one variety to another. First, there must be some communication between the speakers of the two varieties, and second, the receiving variety must have the necessary linguistic characteristics. The reason why particular linguistic characteristics may be required is that, if the particular innovation consists of a change undergone by some particular feature of a linguistic system, only a system which possesses that feature in the first place will be capable of undergoing the diffused innovation.⁵

In short, each variety may be expected to originate some innovations. Innovation in itself is a differentiating factor. That is, each innovation will at first differentiate the originating variety from the other members of the family. However, some innovations will diffuse between varieties. Diffusion has the effect of partially offsetting the differentiating effect of innovation. The amount of diffusion between two varieties will decline as the amount of communication between them declines (other things being equal; that is to say, provided the sociological circumstances of the relation between the two remains constant).⁶ One important influence on the amount of communication will generally be the geographical distance which separates them.

The amount of diffusion may also be expected to decline as the linguistic similarity between the two declines. That is, it will decline in the measure that what we may call their "systemic compatibility" declines.⁷ However, this effect is a very gradual one; there is no single Rubicon-crossing step in the process.

THE NATURE OF THE PREHISTORIC EVENTS TO BE RECONSTRUCTED

As I said, I believe that the wave model permits a more accurate representation of the kinds of events which led from the Oceanic proto-language to the contemporary state of affairs. Of course, we have no direct evidence of those events, but I believe that we would find ourselves in approximate agreement as to their general nature. And I believe that our model of linguistic relationships *should* be compatible with the events by which those patterns of relationship were produced.

I would imagine that most students of the field would agree that, starting from the time when Proto-Oceanic was being spoken in its original homeland, the most likely pattern of spread would have been something like this: From very early some dialect differentiation along geographical lines would have been apparent. However, the differentiation, except in rare cases where communication was completely cut off, would have been in part counterbalanced by the spread of innovations. The pattern of spread would have conformed essentially to the classical wave model.

I would suggest that a typical event in the spread of Oceanic would have been for a group of people, usually from somewhere near the periphery of the area then occupied by Oceanic, to move out and establish a new settlement in some place beyond the old periphery.

Let us imagine this situation. Before the migration, our migrants-to-be live in an area in which some linguistic differentiation has already occurred. The nearest village, both geographically and linguistically, to the village of the migrants-to-be speaks what I will call the "neighbour variety". The actual migrants make up only a part of the population of their village and their speech community. Those who remain behind will be called the "stay-at-homes". Thus, at the time of the migration, the migrant and stay-at-home varieties are identical.⁸

Now let us consider the linguistic situation which develops after the departure of the migrants. We may imagine that from this time on the stay-at-homes will share more innovations with the neighbours (because of their relative accessibility) than they do with their departed next of kin, the migrants. The situation which eventually develops is one which is ambiguous in the family tree framework. The stay-at-homes share features (innovations of the period prior to the departure) with the migrants. These shared innovations suggest that the stay-at-homes subgroup with the migrants. But the stay-at-homes also share innovations (those of the period subsequent to the departure) with the varieties which we are calling the "neighbours". These sharings appear to constitute evidence for a subgroup consisting of the stay-at-homes and the neighbours as opposed to the migrants. In fact, there is no clear basis in what I have told you for determining which subgrouping should be judged to be correct.

As I understand it, the traditional basis for deciding such questions relied on the old assumption that something of particular importance had occurred when what had been a dialect boundary was elevated to the status of a language boundary. That is, it relied on the "Rubicon-crossing" of which we have spoken. The key question would be which boundary, that separating the migrants from the stay-at-homes or that separating the stay-at-homes from the neighbours, first attained the level of a language boundary, as opposed to a dialect boundary. As I have explained, I believe that that is a misconceived question, and therefore I do not want us to have to be concerned with it.

A STRATEGY BASED ON THE WAVE MODEL

Rather than attempt an extended critique of the family tree model here, I will focus instead on four characteristics that I believe are desirable in a strategy for determining the relationships within a language family, and which I will accordingly incorporate in my proposed wave-based model. They are as follows:

(1) We always tacitly assume that the geographical distribution of the languages should correspond in some measure to their genetic linguistic relations. It seems reasonable to expect geography to provide leads and it seems sensible to attempt to make use of them.

I proposed above that the most likely kind of event in the spread of Oceanic languages would have been for a group of people from the then periphery of the Oceanic-speaking area to establish a new settlement somewhere beyond the old periphery. It seems also to be likely that most of these moves would have been short - that the destination would have been an area which was known beforehand by the migrating group and which was not very difficult of access. I propose therefore what I will call the "principle of shortest moves". This principle is that in the absence of evidence to the contrary, it is to be assumed that each new Oceanic settlement was made from the geographically closest Oceanic-speaking place then in existence.

This principle may be generalised to say that, in the absence of evidence to the contrary, the relative linguistic closeness of any two varieties corresponds to the relative geographical proximity of the locations in which they are spoken. The task of the comparative linguist will then become that of identifying cases where varieties are out of place - where the linguistic relations do not conform to their geographical location - and of providing supporting arguments.

(2) We want our strategy to permit us to escape the assumption that language boundaries - that is, the boundaries which separate different languages - are different in kind from dialect boundaries. Therefore, in the approach which I propose one does not need to talk about languages at all. Instead, I will speak of "local varieties". For present purposes I think it will be sufficient to think of a local variety as that which is spoken in one settlement - e.g. a village.

(3) It is desirable to design our strategy so that the allocation of burdens of proof is symmetrical. What I mean is this: in the framework of the family tree model the conventions of argumentation all seem to concern argumentation for one kind of claim - the claim that some subset of the languages of some language family constitute a subgroup. The most often cited statement on the evidence required is that of Karl Brugmann (1884:253) that the proof of a subgrouping consists of a mass of common innovations, phonological, flectional, syntactic, and lexical - a mass so great as to exclude the thought of accident.

There seems to be no comparable discussion of what is required to prove the opposite claim, i.e. that a particular set of languages do not constitute a subgroup. The only strategy available to whoever wants to argue against such a grouping is to try to create doubt about whatever evidence has been adduced or seems likely to be adduced in support of the grouping. We might therefore speak of the kind of evidence appropriate to support a claim that a subgroup exists as *positive* evidence and that appropriate to argue against its existence as *negative* evidence. I call it negative evidence because it can be seen as evidence at all only in relation to, and in reaction against, some adduced or anticipated positive evidence.

Where this difference between the kinds of evidence exists, it is impossible to measure out the burden of proof in equal proportions. That is, it is impossible to design the rules of play in such a way that the requirements of one who would argue for a subgrouping and one who would argue against it are equally rigorous.

In the strategy which I will propose there will be no difference in the kinds of evidence required to support counterclaims as compared to that supporting the original claim.

(4) In Grace 1985 I tried to show that we ordinarily make the tacit assumption that languages which are closely related will be generally similar to one another. As a consequence any perceived lack of similarity between varieties is likely to be construed as *prima facie* evidence that they are not closely related to one another. I tried also to show there that such a tacit assumption rests on another tacit assumption - that the rate at which languages change is approximately constant. I pointed out that there are serious reasons to doubt that that is true.

In view of these considerations, we do not want our strategy to involve any assumption of gradualism, overt or hidden. That is, we assume that rates of change and the cumulative effects of change may differ widely from one instance to another.

APPLICATION OF THE WAVE-BASED STRATEGY

There are a few points which I would like to mention before presenting the results of the application.

First, my own experience convinces me that it is very useful to recognise groupings which consist of chains of most closely related varieties but which may not be subgroups in the strictest sense. That is, they may not be subgroups in the sense that they may not be descended from a proto-language with all of the necessary qualifications. A proper proto-language according to family tree theory must have been a homogeneous entity in the sense of being characterised throughout all of its dialectal range by the same set of linguistic innovations. In addition, it must have been set off by full-fledged language boundaries from any other varieties.

The kind of groups which I am proposing are what I called "subgroups" in the so-called "subgrouping" which I published in 1955 (Grace 1955). Those were intended to be simply chains of varieties such that each chain was completely set off by boundaries from all outside varieties. Thus, it was required that the closest relative of each member variety was another member but not that each member must be more closely related to every other member than to any non-member. I find it very difficult to get along without some such linguistic units, so I will implicitly recognise them here.

Second, we will often encounter gaps in the geographical distribution of the linguistic units (i.e. varieties or chains of varieties) belonging to the family. A gap may be occupied by languages not belonging to the family or it may be unoccupied - that is, it may be uninhabited land or, especially in the ocean area with which we are concerned here, water. Where there are gaps, we may speak of the distribution as being broken up into different linguistic 'islands'. That is, an island will be an area which is occupied throughout by Oceanic linguistic varieties, and which is separated from all other Oceanic varieties by geographical gaps.

The existence of such islands will naturally lead to the question of where each island fits in relation to the others. I expect that in actual practice one of the islands under consideration at a particular time will usually be seen as the matrix island - or 'homeland' - in relation to the others, and that the

problem will turn out to be that of placing the other islands with respect to this matrix island. In accordance with the principle of shortest moves which I described above, I will assume that closeness of relationship corresponds to geographical distance even where such gaps intervene. Thus, the assumption in the absence of contrary evidence will be that even if the closest relative of a particular linguistic unit is not within the same island, it will still be the geographically closest unit - i.e. the unit separated from it by the smallest gap.

The most important kind of question which will arise is whether there are any units which are geographically out of place. That is, are there any cases where varieties or subgroups of different geographical proveniences have come to be juxtaposed by secondary migrations? I will discuss the evidence bearing on questions of this sort below, but first, I want to lay out my tentative hypotheses about Oceanic linguistic prehistory.

HYPOTHESES ABOUT OCEANIC LINGUISTIC PREHISTORY

A few of the hypotheses which are suggested by the proposed wave-based strategy are as follows:

(1) It is hypothesised that the Oceanic homeland was in Irian Jaya (Indonesian New Guinea) in the region including the Sarmi coast and probably the Kumamba Islands. This hypothesis is based on the assumption that the homeland of Oceanic must be the Oceanic area which could have been reached by the shortest move from the area in which the most closely related non-Oceanic Austronesian languages are spoken.

(2) Turning to the areas most remote from this homeland and therefore presumably the last reached by Oceanic speakers, we may, first of all, accept the conclusions of various investigators that the Polynesian settlement results from successive migrations from central Vanuatu to Fiji, thence to Rotuman and also to Polynesia, with the subsequent spread throughout Polynesia and back to the Outliers.

However, the genetic relations of most of the other remote groups have been more problematic. Here are the hypotheses which the principle of shortest moves yields for the languages and groups at the far end of the Melanesian chain, namely the languages of New Caledonia, the Loyalty Islands, Southern Vanuatu, Central and Northern Vanuatu, and the Santa Cruz Islands:

(a) New Caledonia was presumably settled from one of the Loyalty Islands.

(b) Each of the Loyalty Islands was presumably settled from southern or central Vanuatu or from another of the Loyalty Islands (if we were to apply the principle of shortest move strictly, we would have to hypothesise that Mare was settled from Aneityum, and that Lifu was settled from Mare, and Uvea from Lifu. The main island of New Caledonia would then have been settled from Uvea or Lifu).

(c) Southern Vanuatu was presumably settled from central Vanuatu.

(d) Vanuatu was presumably settled from the south-eastern Solomon Islands (probably San Cristobal).

(e) The Santa Cruz Islands (at least as far as the Oceanic languages there are concerned) were presumably settled from San Cristobal or thereabouts.

(3) Perhaps the most conspicuous problem concerns the Oceanic but non-Polynesian languages of Micronesia, which I will just call the 'Micronesian' languages. They presumably reflect a movement from somewhere in the Melanesian chain to somewhere in Micronesia. The closest geographical distance between Micronesian and Melanesian (i.e. Oceanic, but not Micronesian or Polynesian) languages is between the Mortlocks and the St Matthias Islands. However, both Micronesia and Melanesia are large areas, and the distances between them at several different points are very roughly equal. Therefore, it is difficult to make any decision on geographical grounds alone.

Micronesian internal relationships suggest, on the basis of linguistic diversity, that the original settlement of Oceanic speakers in Micronesia was in the eastern part of the area - somewhere in the area from Ponape to Kiribati. The closest points in that area to Melanesian languages are Nauru and Ocean Island. The closest Melanesian languages to Nauru and Ocean Island are in the Solomons (Malaita and Isabel) and the Santa Cruz Islands. The very shortest distance would be that separating Nauru, in Micronesia, and Malaita in Melanesia.

If we were to make these assumptions: (1) that the original Oceanic settlement of Micronesia was most likely in eastern Micronesia, and that the most likely point is the closest point to Melanesia in eastern Micronesia - namely Nauru, and (2) that the most likely point of *origin* is the geographically closest point to Nauru in which a Melanesian language is spoken (that is, if we apply the principle of shortest moves), then the hypothesis would be that people from Malaita (or possibly Isabel, Choiseul, Santa Cruz, or thereabouts) settled on Nauru (or perhaps Ocean Island), and that thence they gradually spread throughout Micronesia.

I should point out at once that the principle of shortest moves does not seem to constitute a very weighty argument where all possible distances are quite long as in this case. However, the hypothesis is made more interesting by the fact that Robert Blust (1984) has recently found some, admittedly skimpy, evidence linking Micronesian with the languages of Malaita.

Most of these hypotheses have not been proposed before, as far as I know. That concerning the position of Fijian and Polynesian is the most noteworthy exception. In addition, Blust has, as I mentioned, suggested a possible connection between Malaitan languages and Micronesian. Finally, one might want to count the hypothesis of the settlement of southern Vanuatu as an exception since I personally have always thought that the southern Vanuatu languages would turn out to be most closely related to other Vanuatu languages.

I am not aware of any previous suggestions about where any of the other groups fit in. What the present approach has done is put forward some relatively specific hypotheses. These are not based on linguistic evidence, of course, but then no one seems to have been able to find any linguistic evidence in those cases. What these hypotheses can claim to be, therefore, are the best guesses available in the *absence* of linguistic evidence. But they might possibly have one further merit; they might provide a specific focus for a future search for linguistic evidence.

Let me say just a few further words about how this approach would reconstruct the history of the spread of Oceanic throughout Melanesia. It suggests that, from the original homeland on the Sarmi coast, Oceanic languages must have spread eastward along the coast of New Guinea. When they reached the Sepik Province of Papua New Guinea, there was presumably a split. One offshoot would have gone north to Wuvulu-Aua and thence eastward through the Admiralties to Emira-Mussau

(St Matthias) and New Ireland. It was presumably this prong of the advance which reached the Solomons - perhaps via Nissan - and thence, as we saw, moved on to eastern Melanesia and Polynesia.

From the Sepik Province, the other prong would have continued eastward along the coast and offshore islands such as the Schoutens. Again there must have been a split on or around the northern coast of Morobe Province with one branch going to New Britain⁹ and the other continuing down the coast to provide the Oceanic languages of Papua.

Some parts of these latter hypotheses have been proposed before, and some supporting evidence for some of them has been published. However, some of the seemingly most significant of them are entirely new, and again, I can only suggest that they represent best guesses in the absence of linguistic evidence and possible foci for a future search for linguistic evidence.

HOW HYPOTHESES OF THIS KIND CAN BE DISPROVED

What kind of argumentation is possible against a claim made in this proposed new game? The possible attacks seem to be of two kinds. One would be to show that I have made mistakes. For example, it is not impossible that I have made mistakes in determining the distances, and that some of the moves which I have proposed are not actually the shortest possible.

The other kind of attack, and this is the most interesting case, would involve disproving one or more of my purported moves by showing that the geographical proximity on which the postulated move was based is the result of a secondary juxtaposition - i.e. that it involves a move other than the minimal sequence of shortest moves needed to account for the present distribution of Oceanic languages. But how may one go about proving this?

First of all, one obvious kind of argument is excluded by one of the assumptions which we have adopted. That assumption is that linguistic change might be greater in some times and places than in others. The main practical consequence of this assumption is that we cannot base a claim against closeness of relationship on the negative grounds of linguistic *dissimilarity* between the linguistic units involved. We are, therefore, left with the following possibilities in a case where a unit A is supposed on the basis of its geographical proximity to have a unit B as its closest relative:

- (1) Possibly upon closer examination, A can be shown to share innovations exclusively with B or with B and a few of B's close relatives. If so, that is evidence that the geographically-based hypothesis was right, after all.
- (2) Possibly A can be shown to share innovations exclusively with other units geographically and linguistically remote from B. If so, that is evidence that some unforeseen migration has brought A secondarily into this geographical juxtaposition with B. In this case, A's true next-of-kin will have been identified.
- (3) Possibly A can be shown to have retained features of an earlier proto-language which have been *replaced* in some of the innovations which are shared exclusively by B and some of B's closest relatives. If so, that is evidence at least that B's relation to A is more remote than its relation to the languages which share those innovations. This would be at least compatible with A's location being a secondary juxtaposition.

CONCLUDING REMARKS

This scheme is being put before you as a serious proposal. Of course, it will require much further testing before a final evaluation of its usefulness can be made, but its underlying assumptions appear to be intuitively sound and it seems likely to provide valuable research leads. Furthermore, the application of it which I have attempted here has generated a whole series of more or less novel hypotheses.

I believe that it offers two important advantages over our present strategy. The first is that it provides an alternative to the family tree with the latter's reliance on the concept of the subgroup. The difficulty with the subgroup concept is that the kind of discrete and homogeneous entities that are required as proto-languages for the subgroups must surely have occurred only sporadically in the history of any language family. Any strategy for determining the genetic relations among a group of languages which is prevented from recognising any relationships except those which can be expressed in terms of such subgroups seems destined to provide very unsatisfactory results.

The other advantage which it offers is that it makes reasonable hypotheses about genetic relationships available for use even before all of the linguistic evidence has been compiled. These hypotheses should be of value to the culture historian as well as to the linguist. I say that the hypotheses are reasonable because I believe that in most cases it will turn out to be true that the nearest relative of a particular variety will be the variety which is geographically closest to it, as the theory holds. Moreover, I believe that in the majority of cases where the closest relative turns out *not* to be the nearest geographical neighbour, our hypothesis will prove still not to have been wrong by very much. That is, I predict that in most such cases the closest relative will still be relatively near, and that the nearest neighbour will turn out to be a fairly close relative. In short, I predict that the hypotheses will usually be right, and that when they are wrong, the error will most often amount to no more than a minor imprecision.

I believe that it deserves the serious consideration of both culture historians and linguists.

NOTES

1. This is a slightly revised version of the historical linguistics keynote address delivered on August 16, 1984, to the Fourth International Conference on Austronesian Linguistics, Suva, Fiji.
2. In fact, the 'network-breaking' model of linguistic differentiation proposed in Pawley and Green 1984, comes quite close to that proposed here.
3. The idea of defining closer relationship as a relation holding between individual languages independent of claims about subgrouping as such probably was first suggested to me by Isidore Dyen. Dyen (1953:580) wrote:

In contradistinction to Brugmann's usage, closer relationship can be used as a broad term to include both membership in the same dialect area of the proto-language as well as membership in the same subgroup, where a subgroup implies a private proto-

language which was one of the languages resulting from the differentiation of the proto-language of the whole family. In this sense a closer relationship is indicated by any exclusively shared innovation.

4. Actually, there is one qualification which must be made in the case of language boundaries. It is generally acknowledged that lexical items may diffuse across language boundaries. However, it is customary to think of the lexicon of a language as more nearly an accidental accumulation than as contributing to the essential character of the language. Therefore, this qualification is generally not regarded as a serious one.

I think that this sharp distinction between lexicon and grammar (or whatever the other parts of language dealt with in linguistic descriptions are to be called) is a misleading exaggeration. See my discussion of the process of differentiation just below.

5. I presume that this requirement of systemic compatibility is the source of the idea that only lexical items can be borrowed from one language to another.
6. Not nearly enough is known about this, but as John Wolff has reminded me, there are as many different kinds of language contact situations as there are kinds of possible sociological relations between the languages or varieties in contact. Each different kind of contact situation is likely to manifest a different pattern of diffusion. In the model I am describing I assume the sociological relations to remain unchanged.
7. The suggestion that the amount of diffusion will be in some inverse relationship to the degree of "systemic compatibility" again assumes everything else to be equal. Furthermore, it should not be understood as claiming that very great structural differences cannot be levelled given favourable circumstances and sufficient time.
8. The same problem would arise if, instead of being linguistically identical at the beginning, the varieties spoken by the migrants and the stay-at-homes were simply more closely related to each other than to that of the neighbours, as long as there was enough systemic compatibility between stay-at-home and neighbour varieties to permit fairly free diffusion of subsequent innovations.
9. The Tolai area of New Britain was surely settled by the Oceanic prong which passed through New Ireland. Therefore, Oceanic languages entered New Britain from both ends. My guess at this point would be that all present-day Oceanic languages in New Britain except the varieties of Tolai are traceable back to the settlement from the Morobe coast.

REFERENCES

BLUST, Robert

- 1984 Malaita-Micronesian: an Eastern Oceanic subgroup? *Journal of the Polynesian Society* 93:99-140.

BRUGMANN, Karl

- 1884 Zur Frage nach den Verwandtschaftsverhältnissen der indogermanischen Sprachen. *Internationale Zeitschrift für allgemeine Sprachwissenschaft* 1:226-256.

DYEN, Isidore

- 1953 Review of Otto Christian Dahl: *Malgache et maanjaan: Une comparaison linguistique*. *Language* 29:577-590.
- 1965 *A lexicostatistical classification of the Austronesian languages*. *Indiana University Publications in Anthropology and Linguistics*, Memoir 19. (Supplement to *International Journal of American Linguistics*).

GRACE, George W.

- 1955 Subgrouping of Malayo-Polynesian: a report of tentative findings. *American Anthropologist* 57:337-339.
- 1985 Oceanic subgrouping: retrospect and prospect. In Andrew Pawley, and Lois Carrington, eds *Austronesian linguistics at the 15th Pacific Science Congress*, 1-18. PL, C-88.

PAWLEY, Andrew and Roger C. GREEN

- 1984 The Proto-Oceanic language community. *Journal of Pacific History* 19:123-146.

ON THE NATURE OF SUBGROUPING ARGUMENTS

S.P. Harrison

A warrant is a theoretical premise governing arguments in scientific discourse, a premise that licenses the move from data to conclusions in a particular domain. The warrant can be viewed as the major premise in an argument whose minor premise is a datum or data and whose conclusion is an inference from that datum or data. The warrant, then, is a means of evaluating the strength of particular data as evidence for a particular conclusion. An acceptable warrant need not, in any strong sense, be one that has been independently verified. Within a theory it may have axiomatic status, the status of a rule by which the game is played. In order to be judged minimally acceptable, however, it seems to me that a warrant:

- i) must be consistent with other premises on the theory, unless appropriately qualified
- ii) must be relevant to the data used
- and iii) must entail a deductive argument form, in which the conclusion is valid.

It is my contention that the warrants currently employed, explicitly or, more often, implicitly, in the subgrouping enterprise fail one or more of these criteria of consistency, relevance, or validity.

I will begin* by distinguishing two types of genetic argument in comparative historical linguistics. An *in vacuo genetic argument* is an argument that the members (>2) of some set D of languages do or do not share a common ancestor.¹ A *subgrouping argument* is an argument that the members (>2) of some proper subset S of a set D of genetically related languages share a common ancestor and that no members of the complement of S in D are descended from that ancestor.

The major warrant for *in vacuo* genetic arguments is given by Anttila (1972: 302) in the following form:

- (1) If two or more languages share a feature which is unlikely to have occurred spontaneously in each of them [or been borrowed into one or more of them - SPH],² this feature must have arisen once only, when the languages were one and the same.

The property 'share a feature' requires interpretation. In what is termed the *comparative method*, the shared features (or similarities) are represented as relations (or correspondences) among component segments of signantia (with similar signata) in the languages compared. The method requires that these relations (*correspondences*) be identical for all the relevant signantia compared. This requirement is usually termed the *regularity principle* (see further below).

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 13-21. *Pacific Linguistics*, C-94, 1986.

© S.P. Harrison

13

The above warrant suggests a classic argument form, a dilemma or disjunctive syllogism, by which, in this case, one demonstrates genetic relatedness by successively ruling out borrowing, chance, and iconism as accounts of shared similarities. In principle, each of the steps in this argument requires its own (minor) warrant. Iconism is ruled out a priori, however, by filtering the data for comparison, such that only similarities among signantia related symbolically (by an arbitrary convention) to their signata count as evidence. If one, in addition, wanted to rule out chance at this stage in the argument, as it seems to me some comparativists do, a minor warrant of the following form would be required:

- (2) Similarity (regular correspondence) among symbolic signantia is unlikely to be due to chance.

The identification of borrowings is a less trivial matter. Except in cases of massive borrowing (such as reported in Biggs 1965), many borrowings will show up as irregular correspondences. This observation can be incorporated into a revised version of (2) above:

- (3) Regular correspondence among symbolic signantia is unlikely to be due to either chance or borrowing.

Having thus filtered out iconism, chance, and borrowing as sources of similarity, one can employ the remaining data as evidence that the languages compared are genetically related. (Other considerations bearing on the cognacy (genetic relatedness) of signantia, for example, similarity of signata, or number of component segments of the signantia, will not be considered here.)

There seems to be an assumption, widely held, that essentially the same major warrant and, by implication, the same argument form, governs both in vacuo genetic and subgrouping arguments. One might note, for example, that Anttila offers his version ((1) above) of the 'genetic warrant' as a subgrouping warrant, noting subsequently that "the establishment of a language family uses the same principle ..." (1972:302).

The subgrouping warrant must differ from the genetic warrant in at least one respect, however; in attending to the possibility that observed similarities among the members of a putative subgroup are not due to retention from a higher order (superset) ancestor. This requirement is usually stated in such terms as the following:

- (4) The conclusion that some set of languages form a subgroup is licensed by the demonstration that the members of that set uniquely or exclusively share one or more innovations attributable to the most immediate ancestor of the members of that set.

It might appear that the subgrouping warrant can be derived from the genetic warrant simply by appending an additional disjunct of the form:

- (5) ... or [that the similarity not be] a feature retained from a more remote ancestor ...

as a further necessary condition, together with an appropriate minor warrant. The original genetic warrant and the derived subgrouping warrant do differ in one significant respect, however. The former is a warrant to *conclude* that the members of some set of languages are genetically related, while, in the latter, the genetic relatedness of the languages compared is a *premise*, explicit in the characterisation of a subgrouping argument given above and implicit in the phrase 'more remote ancestor'.

This difference is crucial to the nature of the data relevant to a subgrouping argument. One might observe first that subgrouping arguments are subsequent to in vacuo genetic arguments. That is, one has already concluded that the members of some set of languages are genetically related, through an in vacuo genetic argument, before one attempts to construct subgrouping arguments. It follows, then, that the data available for a subgrouping argument is of two sorts:

- i) the data available for the antecedent genetic argument
- and ii) reconstructions³ for an ancestor of the members of a set of languages, some proper subset of which, it will be argued, constitutes a subgroup.

The correspondence between comparable elements in these two domains (that is, between the reconstructions and the primary data) is usually represented as a set of functions called *changes*, which are themselves interpreted as (sequences of) historical events. It is these functions, the changes, and not simply observed similarities among signantia, that constitute the data for a subgrouping argument. If one observes that two or more members of the set of languages compared share identical changes, one then asks whether these identical changes must be interpreted as the same historical event. If they must be so interpreted, one is licensed to conclude that the languages in question form a subgroup. Thus, what is required for subgrouping is a warrant to conclude that identical changes are a single historical event.⁴

It seems to me reasonable and perhaps advisable to construct a subgrouping warrant that is formally similar to the genetic warrant; that is, as a series of disjuncts, each of which is to be ruled out in turn (in a subargument) such that only the last, that the changes are a single historical event, remains. These disjuncts themselves might be viewed as parallel to those of the genetic warrant; that is, that the identical changes are the result of:

- i) diffusion after the diversification of some common ancestor has begun, or
- ii) accidental parallel separate development, or
- iii) motivated parallel separate development, or
- iv) a single historical event

corresponding to borrowing, chance, iconism, and genetic relatedness, respectively. There have, to my knowledge, been no explicit attempts to construct a major subgrouping warrant along these or similar lines. (In a subgrouping argument one must also recognise the possibility that the higher order reconstructions, from which the changes were inferred, might be incorrect. If this were the case, then the putative identical changes to be evaluated in terms of i)-iv) above are not changes at all, but retentions from a higher order ancestor and, as such, do not constitute data for a subgrouping argument. This issue is considered further below.)

There are, I believe, two distinct premises which, though not in fact subgrouping warrants, do govern the present conduct of the subgrouping enterprise.⁵ The first I will term the *fact of change premise*:

- (6) Given the theory of change T, the changes $c_1 \dots c_n$ form a hierarchy of naturalness, predictability, explicability, or frequency.

and the second, the *act of change premise*:

- (7) The incidence of any change is unnecessary and, therefore, unexpected and remarkable (though the nature of the change may not be odd of itself).

These premises are not subgrouping warrants because they do not directly license the conclusion that a particular shared change is a single historical event. In order to construct a warrant from the fact of change premise one would have to accept that the more 'unusual' a change, the more likely it is to have occurred only once (with respect to the disjunctive conditions above, the less likely it is to have diffused, occurred by chance, or, by definition, been motivated). Such a warrant, it seems to me, corresponds to the symbolic data filter in the comparative method and thus, rules out functional motivation and chance as accounts of shared changes.

The problem with the premise (6), and with any warrant derived from it, is that there is no theory T in which a suitable hierarchy of changes is defined. In fact, the major thrust of contemporary historical linguistic theory has been in precisely the opposite direction, to the effect that all linguistic changes (in this context, types of change) are in some sense motivated, either articulatorily, perceptually, or indexically.⁶

Consider, for example, a subclass of the class of what Pawley (1974) terms *replacement innovations*, changes whereby "some feature *x is replaced by a new feature y ($\neq 0$ - SPH), which is in some sense its continuation or functional equivalent". The subclass I have in mind involves cases of the replacement of one lexical item by another, innovations which Pawley (1974:5) feels "can only rarely be established". Were such an innovation established, however, one need not interpret it as 'odd' in the intended sense, since such replacements are not unmotivated. The result of such a change is, in the prototypical case, the loss of the replaced signans and a change in the signatum of the replacing signans. Crucially, the replacing signans may, in principle, have any form whatsoever, but not so the replacing signatum. To my mind, it is difficult to conceive of replacing signatum that is neither metaphorically nor metonymically related to the replaced signatum. Thus, though the act of change *might* be unmotivated in such cases, the fact of change is not. *Mutatis mutandis*, the same ought to hold for replacements in other domains. It is inconsistent to maintain the fact of change premise, as the basis for a subgrouping warrant, and at the same time maintain the thesis that change types are motivated.

The act of change premise (7) is less controversial since, I believe, few historical linguists would want to maintain that there are sufficient conditions for the act (incidence) of change. It is difficult to formulate this premise precisely though, given that agreement is far from complete among historical linguists regarding just how surprised one should be when a particular act of change takes place.⁷ The premise is crucial, however, to the methodology of 'realistic' reconstruction (see note 3), as the basis for the heuristic favouring that reconstruction which requires the fewest number of individual acts of change.

If one felt that what might be termed the *singularity theorem*:

- (8) Since all acts of change are remarkable, all formally identical changes in a given language are likely to be the same historical event

can be derived from the act of change premise, it is possible to relate that premise to the regularity principle governing the interpretation of sound change. Though the regularity principle was formulated above as an identity condition on correspondences, it can be, and frequently is interpreted as a premise regarding acts of sound change:

- (9) The act of (sound) change is a single event, affecting all tokens of a given type for all speakers in a well-defined speech community simultaneously.

While such an interpretation of the regularity principle has long been regarded as demonstrably false, this fact detracts only minimally from its utility in the context of the comparative method, since it is also well known that a significant number of observed correspondences are regular, no matter how many separate acts of change, in each of the languages compared, were required to establish this regularity.

The act of change premise is thus crucial to the methodology of (in vacuo) genetic comparison in two respects; in realistic reconstruction and, through the singularity theorem, in an idealised interpretation of the effect of change on signantia in a single language. The singularity theorem can also be modified as a major subgrouping warrant:

- (10) Since all acts of change are remarkable, if two or more languages share the same change then this change is likely to have occurred only once, in the most immediate ancestor of the languages sharing this change.

This *strong act of change subgrouping warrant* in effect holds that any shared change is evidence for subgrouping. Some of the more outlandish subgrouping claims licensed under this warrant might be rejected by invoking the geographic proximity warrant but, even with this qualification, few comparativists would admit to maintaining the warrant in this strong form.

The problem lies with the act of change premise (7), on which the warrant (10) is based. This premise is essentially a claim that acts of change are not motivated (iconically, for example) within the linguistic system itself. One can maintain this premise without at the same time being committed to the view that two or more languages cannot come to share changes either:

- i) as a result of separate accidents (i.e. chance)
- or ii) as a result of contact stimulus diffusion (i.e. borrowing).

(Indeed, chance, in a broad sense, and diffusion appear to be the only two possible interpretations for shared acts of change under the act of change premise.) Using the act of change premise, a valid subgrouping argument must rule out both diffusion and independent accident (after the members of the putative subgroup had begun to diverge) in order to license the conclusion that the shared change can be attributed to a single historical accident/event. One therefore requires subwarrants for each of these subarguments.

One might first dispose of what I will term the *temporal sequence subwarrant*:

- (11) If it can be demonstrated that a change c_b , shared by L_1 and L_2 , followed a change c_a found only in L_1 , then c_b developed subsequent to the divergence of L_1 and L_2 .

Though innocuous, the subwarrant (11) is *not* a subgrouping warrant of the desired sort. It licenses the conclusion that a shared change is not a single historical act, but the result of either independent accidents or of stimulus diffusion. What is required of a subgrouping warrant is licence to conclude that the observation of shared change is not interpretable in either of these ways. The subwarrant (11) is useful only in arguments that the members of some set of languages are *not* a subgroup.

It seems to me that stimulus diffusion can never be ruled out, except on a priori or on geographical grounds. Though I believe that some comparativists, until very recently, have adopted an aprioristic stance with regard to the possibility of the stimulus diffusion of a change after diversification has begun, there is ample evidence that such changes do take place. The same can be said of parallel separate development (independent accident). On these grounds, one is safe in rejecting the strong form (10) of the act of change subgrouping warrant.

As already noted, most comparativists would in fact reject that warrant in favour of one or both of the following weaker versions:

- (12) Since the act of change is unmotivated/accidental, the more languages share a change, the more likely it is that that change is a single historical event.
- (13) Since the act of change is unmotivated/accidental, the more changes are shared by some set of languages, the more likely it is that the members of that set form a subgroup.

Subgrouping warrant (12), which I will term the *weak act of change (token) subgrouping warrant*, is, I think, seldom recognised. Though, like the temporal sequence subwarrant (11), it is innocuous, it is of questionable utility since it asserts only that evidence for subgroups with more members is stronger than evidence for subgroups with fewer members. It is, in some sense, irrelevant to the data since it provides only a partial function on the domain of recognised subgrouping evidence. Subgrouping warrant (13), which I will term the *weak act of change (type) subgrouping warrant* is also unacceptable, because it requires some metric by which relative quantity of shared changes can be interpreted in terms of the *family tree model*. In the absence of such a metric it seems to me that the model of genetic relatedness most appropriate to the interpretation of such quantitative observations is the *wave model*, in which case the issue of subgrouping becomes moot.⁸

In summary, as presently conducted the subgrouping enterprise is based, implicitly at least, on one or more of the following warrants:

- i) the geographical proximity warrant
- ii) the fact of change warrant
- iii) the strong act of change warrant
- iv) the temporal sequence warrant
- v) the weak act of change (token) warrant
- vi) the weak act of change (type) warrant.

It has been my contention that no one of these warrants, nor any combination thereof, provides an acceptable basis for subgrouping. The geographical proximity warrant is uncontroversial but irrelevant, in the sense that it provides no basis for assessing linguistic evidence. The fact of change warrant lacks a necessary theoretical basis and, in fact, is inconsistent with most current theory. The strong act of change warrant does not provide for a deductive argument form in that it does not rule out, nor does it seem capable of ruling out either stimulus diffusion or separate chance actuation as accounts of shared changes. The temporal sequence warrant licenses only negative subgrouping conclusions. The weak act of change (token) warrant provides only a partial function on the domain of subgrouping evidence. The weak act of change (type) warrant is inconsistent with the family tree model. I must, then, conclude that the subgrouping enterprise, as currently conducted, is incoherent because it is being conducted in the absence of any acceptable warrant.

This assessment of the status of subgrouping arguments, if at all valid, cannot but be disturbing to comparative historical linguists, who might then conclude that the foundations of the discipline are under attack. That conclusion does not necessarily follow. Though subgrouping is a traditional concern of comparative historical linguistics, it does not follow that the discipline stands or falls on that basis. Such would only be the case if the determination of subgroups were methodologically prior to other goals of comparative historical linguistics.⁹

These other goals are:

- i) the determination of in vacuo genetic relationships
- ii) the construction of a theory of change
- iii) the reconstruction of unattested language states.

The first of these goals, as argued above, is independent of and prior to subgrouping considerations. The second is also independent of subgrouping considerations, since it is concerned with the interpretation (explanation/explication) of changes. Subgrouping becomes relevant to the construction of a theory of change indirectly however, if changes attested through reconstruction are admitted as data for a theory of change and to the extent that reconstruction is dependent on subgrouping considerations. It is well-established that this last condition does in fact hold.

If one accepts the heuristic favouring that reconstruction which minimises the number of separate acts of change, then subgrouping premises become crucial, since they affect that number. Given n related languages among which, let us say, two corresponding forms a and b are distributed, one reconstructs either $*a$ or $*b$ (except under typological conditions pointing to some $*c$, which will be ignored here) depending on which is attested in the greater number of languages. If no internal subgrouping is assumed, one simply counts. If the set of n languages is subgrouped on the basis of the distribution of a/b , subgrouping is also irrelevant to the choice of $*a$ or $*b$, since both are now evenly distributed. Only typological considerations or some theory of change can decide the issue in such cases. If the set of n languages is subgrouped on some basis other than the distribution of a/b , such that both a and b are represented in members of all highest order subgroups, the subgrouping premise is similarly irrelevant to the choice between $*a$ and $*b$. Only if a and b are not represented in all highest order subgroups is that subgrouping relevant and, indeed, crucial to the choice between $*a$ and $*b$. Such cases, unfortunately, are common. Even more unfortunate is the fact that they admit of no resolution, since the shared change(s), whatever it/they may be, on which one bases the subgrouping premise being brought to bear on the $*a/*b$ reconstruction problem, is/are, in principle, no better evidence for subgrouping than is the distribution of a/b itself. Appeals to subgrouping in making reconstructions thus dissolve into circularity. Unless other criteria can be brought to bear, such reconstruction problems must go unresolved. But, of course, one need be no more disturbed by this fact than by the fact that reconstruction can be problematic in cases where subgrouping premises are irrelevant.

In conclusion, let me stress that I am not claiming that there are no subgroups, as defined, only that there are no acceptable subgrouping arguments.¹⁰ One cannot hope to justify a subgrouping hypothesis through the search for uniquely shared innovations if one has no warrant to find them. If there is to be a solution to this problem, it must be in the construction of a new, and acceptable warrant for subgrouping arguments. Of the two premises relevant to linguistic data, I see no hope whatsoever in the act of change premise. As it

stands, it does not provide an acceptable basis for the construction of a subgrouping warrant. Nor, of course, does its converse - that acts of change are motivated and expected. There may be more hope in the fact of change premise, but only insofar as the investigation of changes leads to the construction of a theory T in which changes are ranked in a hierarchy of 'naturalness'. (It follows, of course, that the less natural a change is, the fewer instances of that change will be encountered; in short, good subgrouping evidence under such a warrant will be hard to find.) If there are other directions that might be pursued in the search for an acceptable subgrouping warrant, at present they are not at all obvious to me.

NOTES

*The paper first appeared in *Working Papers in Linguistics, University of Hawaii* 15/2:79-91.

1. I define 'genetically related' in terms of the notion 'common ancestor', rather than in terms of the often substituted notion 'period of common history'. The latter properly characterises 'historically related', which includes 'genetically related' but could equally apply to relationship through diffusion/contact/borrowing. The more restricted concept 'genetically related' can, I believe, be explicated in such terms as 'having similarities attributable to retention from an antecedent linguistic tradition'.
2. Anttila's statement of the warrant does not contain this last condition, a defect he later acknowledges. I take his 'spontaneously' to be equivalent to Greenberg's (1953:268ff) 'due to convergence', subsuming both *chance* (convergence through limited possibilities) and *iconism* (convergence through similarity of function).
3. The reconstructions may be either 'realistic' or 'formulaic', the former given phonetic content and the latter not.
4. Similar observations, though with rather different conclusions, are made in Greenberg 1957:48ff.
5. One might note the *geographical proximity warrant*, which many comparativists would not openly sanction:

The more/less geographically contiguous are the members of a set of genetically related languages, the more/less likely it is that they form a subgroup.

This warrant is given a central role in the novel approach to subgrouping suggested in Grace (in this volume).

6. Note that the various implicational hierarchies that have been proposed in recent work govern not the naturalness of changes, but their progress through time. Changes that violate these hierarchies would not be considered to be unnatural changes. Rather, their incidence would suggest deficiencies in the theory. One could, of course, envision an alternative theory, of the type T above, in which the naturalness of changes is evaluated, and in terms of which a subgrouping warrant could be constructed.

7. Given differing views regarding the status of the fact of change premise or regarding the strength of necessary conditions for change.
8. Note that the supposed correspondence between a family tree diagram and a non-overlapping wave diagram is irrelevant to the interpretation of the quantity of shared changes. Such a correspondence could only be licensed by some version of the strong act of change warrant, rejected earlier. The same is true, I believe, of the proposal (elaborated in Krishnamurti et al 1983) for subgrouping based on lexical diffusion.
9. There is a non-linguistic justification for subgrouping, by which family trees are brought to bear as evidence for prehistory. While such altruism is commendable, it does have limits, an obvious one being the status of the subgrouping conclusions. If such conclusions cannot be trusted, they are of little value to anyone.
10. It seems to me that an attempt to construct such an argument might prove a valuable exercise. More specifically, I have in mind a weak version of that argument, to the effect that one should not expect to find clearly demarcated subgroups with all n-membered sets of genetically related languages. The basis of the argument might be the existence of 'dialect chains', within which no subgrouping is expected, together with the hypothesis that the most immediate common ancestor of a given dialect chain was itself a point in a dialect chain. I will not develop this argument further here, however. (The germ of this argument is, I think, implicit in observations made in Grace 1962.)

REFERENCES

- ANTTILA, R.
 1972 *An introduction to historical and comparative linguistics*. New York: Macmillan.
- BIGGS, B.
 1965 Direct and indirect inheritance in Rotuman. *Lingua* 13:373-405.
- GRACE, G.W.
 1962 Problems in Oceanic linguistic subgrouping. *Transactions of the New York Academy of Sciences, Series II* 24/6:684-689.
 1986 Further thoughts on Oceanic subgrouping. In this volume, pp.1-12.
- GREENBERG, J.H.
 1953 Historical linguistics and unwritten languages. In A.L. Kroeber, ed. *Anthropology today*, 265-286. Chicago: University of Chicago Press.
 1957 *Essays in linguistics*. Chicago: University of Chicago Press.
- KRISHNAMURTI, Bh., L. MOSES and D.G. DANFORTH
 1983 Unchanged cognates as a criterion in linguistic subgrouping. *Language* 59/2:541-568.
- PAWLEY, A.K.
 1977 On redefining 'Eastern Oceanic'. MS.

PROTO-AUSTRONESIAN LATERALS AND NASALS

Isidore Dyen and Shigeru Tsuchida

Tsuchida (1976:139-143) suggested that there was enough evidence to support the reconstruction of two different Proto-Austronesian phonemes that he labelled *N and *L. The novelty in his proposal was in the hypothesis of *L, for *N - under the symbol *n₂ - had already been suggested by Ogawa and Asai (1935:6f) and is well supported.

Dahl (1981:101ff) attempted to show that the different reflexes assigned to *N and *L were complementarily distributed, contrary cases being dismissed as either possibly due to error or to assimilative or analogical changes. In his view the correspondences assigned to *L appeared only in initial position whereas those for *N appeared in medial and final positions. Dahl prefers the symbol *ʎ for the single proto-phoneme, but we will use *Da.L for it as a mechanical substitution without affecting his phonetic interpretation which seems to be that *Da.L was a voiceless lateral (Dahl 1976:75).

An important point might be at stake in the issue of complementation, for if *N and Tsuchida's *L are different phonemes, and if they are not distinguished by any Formosan language, their merger could be interpreted as a common innovation supporting the hypothesis of a Proto-Formosan.

For the purposes of the following discussion it is convenient to speak of *N reflexes and *l reflexes. In effect we will mechanically substitute *l for Tsuchida's term *L. At the same time we will mechanically substitute *L for Tsuchida's term *ʎ. To avoid compounding confusion we will label the old *L as *Ts.L and the old *ʎ as *Ts.l.

The substitution of *l for *Ts.L and *L for *Ts.l seems obligatory if *N and *l are different phonemes. None of the Formosan languages in Tsuchida 1976 show different reflexes for *N and *l. He distinguished these proto-phonemes by the reflexes that appear in the non-Formosan languages. Non-Formosan languages appear to reflect *N with a nasal and *l with a lateral. Furthermore many Formosan languages offer a clear [ʎ], sometimes described as being slightly palatalised. There is thus reason to consider whether *Ts.L (= *l) was not phonetically [ʎ], a voiced lateral.

There is some evidence that can be cited in favour of regarding *l as having been voiceless. Saaroa everywhere and one Ami dialect in medial and final positions (Valangaw) clearly offer voiceless lateral reflexes for *l, and Tsou exhibits h. On the other hand the Sakizaya dialect of Ami presents a voiced stop throughout, whereas northern dialects of Ami have voiced apical stops and/or spirants in initial and medial positions and a voiceless spirant in final position. Central and southern Ami dialects have lateral fricatives which are voiced in

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 23-31. *Pacific Linguistics*, C-94, 1986.

© Isidore Dyen and Shigeru Tsuchida

23

initial and medial positions and voiceless in final position. Valangaw has a lateral fricative which is voiced in initial position and voiceless when medial or final (see above). Saaroa, Tsou, and Ami have merged *l with *N, now by some regarded to have been a voiceless nasal. Bunun likewise shows a merger of *l and *N, but the outcome is n, a voiced apical nasal. With the exception of Kavalan all of the other languages exhibit a merger in the reflexes of *l and *N. There is however a difference of opinion in that Tsuchida describes the l of Rukai so-called dialects as 'a voiceless lateral fricative' (1976:112) and Li has described it implicitly as voiced in a publication (1977:5) and explicitly as voiced in a personal communication. [Tsuchida has not yet had an opportunity to re-examine the matter.]

For *L (= *Ts.l) many Formosan languages present a lateral flap. These languages are Kanakanabu, Saaroa, Budai and Mantaauran Rukai, Paiwan, Puyuma, Ami, Saisiyat, and Pazeh. Those that do not, offer reflexes that are compatible without difficulty with a hypothesis that they had a flap articulation earlier on. These are Tsou, Maga Rukai r [ɹ], Sedeq, Thao r [r], Tanan Rukai, Bunun Ø, and Atayal y, Ø. One of the principal writers on the Formosan languages, Paul J-k. Li, has elected to indicate the flap by L, a convenient solution. There is thus evidence that could be used to support the hypothesis that *Ts.l (= *L) was a voiced lateral flap. If *l (*Ts.L) was a voiced lateral, then the interpretation of *L (*Ts.l) as a flap seems to be the simplest hypothesis. Examples of reconstructions containing *L are the following:

PAN teLuH₂, Sed turu?, SaiTa tuLu? (A: u/e), Paz turu? (A: u/e), Pai čeLu, RukTa tuLú, RukBd túLu, RukMg túru, RukTo tuú, RukMn tuLu (all Rukai dialects A: u/e), Kan tuúLu?, Sar tuulu?, Tso turu (all Tsouic A: u/e), Ami tuLu? (A: u/e), BunNC tau, BunS tau?, Tha tu:ru? (A: u/e), Kvl u-tuLu (A: u/e) *three*, Puy ta-teLu? *three persons*, To tolu *three*.

PAN ZaLan, AtyMx raan (women's speech), SaiTa raLan, Paz daran, Pai jaLan, Puy da-daLan, RukTa ka-daLan-a(ne), RukBd ka-daa-daLán-ane, RukMg da-drán+, RukTo da-daáne, Kan caáne?, Sar saLa?a?, Tso cron+, AmiSk zazan (A: z/L), AmiNCS LaLan (A: L/r), BunNCS daan, Tha sa:ran, Kvl Lazan (M: L-z/z-L), To hala *road*, *path*.

PAN Lima?, Sed rima?, Pai Lima, Puy Lima?, RukTa Limá, RukBd Líma, RukMg ríma, RukTo imá, RukMn Lima, Kan Liíma?, Sar ku-Lima?, Tso rimo, Ami Lima?, BunNC hima?, BunS ?ima?, Tha ri:ma?, Kvl u-Lima, Sm lima *five*.

PAN Laŋaw, SaiTa Laŋaw, Paz raŋaw, Pai La-Laŋaw, Puy a-ŋaLaw (M: ŋ-L/L-ŋ), Ami La-Laŋaw, Tha ránaw, Kvl Laŋaw, To laŋo *fly*, RukTa a-La-Laŋáw, RukBd a-La-Láŋaw, RukMg a-ŋároo (M: ŋ-r/r-ŋ), RukTo a-ŋáaw (M: ŋ-*L/*L-ŋ), RukMn a-ŋaLau (M: ŋ-L/L-ŋ) *bluebottle*, Kan taa-ŋaLáu? *gnat*.

PHN biq₁₃eL, AtyMb biqiy (A: i/e), SedTn biqir (A: i/e), PaiTamali biqel ([?] b for anticipated v), Kan vi?íLi?, Sar vi?iLi?, Tso f?iri (all Tsouic A: i/e), BunNC biqi, BunS bihi? (all Bunun A: i/e), Ilk biqel, Ifg bi:ol *goitre*.

The evidence for *N in medial and final positions seems to be indisputable. The Formosan languages offer the same reflexes as for *l and the non-Formosan languages offer the same reflexes as for *n:

PAN CuNuH₁, Kan -cúnu?, Sar -culu?, Tso -cuhu, RukBd -cúlu, Pai culu, Ami -tuluh, BunS -tunu?, BunNC -tunu, Sai -suloh *to roast over a fire*, RukMn culu-*a smell of burning feathers*, Mal tunu *to burn*, To tunu *to cook on an open fire*.

PAN D₂aNum, Kan canúmu?, Sar salumu?, Tso chumu, Pai zalum, Puy zanum (A: n/l), Ami nanum (A: n/l, n/r), Bun danum, Tha sa:öum, Kvl zanum, Sai ralum, Paz dalam *water*, To lanu *to wash or rinse in fresh water*.

PAN buLaN, Kan vu^hne?, Sar vuLale?, Tso froh⁺, Ami vuLal, Bun buan, Kvl buLan, Mal bulan, Fi vula *moon*.

PAN q₂uZaN, Kan ?uc^hne?, Sar usale? *rain*, Tso m-⁺ich⁺ *to rain*, RukBd ^udale, RukMg ud^hl⁺, Pai qujal, Puy Hudal, Ami quLal, Buns hudan, BunNC qudan, Tha qusa^o, Kvl ?uzan, Sai ?^ä-?oral, Aty qwal-ax, Mal (h)ujan, To ?uha *rain*.

The evidence for *l in medial position appears to be sufficient:

PHN bilaŋ, Kvl, Mal bilaŋ, Itb -vilaŋ, Ivt -vidaŋ *to count*.

PHN bulaw(-an), RukTa bulav^á, RukBd bulávane, RukMg blávni, blávne (Tsuchida), RukTo bulávane *copper*, Pai vulavan *copper, brass*, Puy vulawan *brass*, Ami vulawan *gold, silver*, Hlg, Ilk bula:wan, Ngj bulaw *gold*, Tag, Bik, SL bulaw *red*. Under this hypothesis, Ivt vuhawan *gold* would have to be a loanword. However for another instance of Batanic h/x for an etymon regarded before as having *l, see *[tT]aNam below.

PHN bulay, Kan vun^ái?, Sar vuli?i?, Itb vulay *snake*.

PHN gelaŋ, Ami kalaŋ (A: a/⁺), Mal g⁺laŋ *bracelet*.

PHN kaliC, Puy kaliT *fur*, RukTo kalíci *hide, leather*, Kmb kalittu *skin, hide*.

PAN kili?, Kan níki-niki? (M: n-k/k-l), Sar líi-likí? (M: l-k/k-l), Tag, Bik kili-kili, Ilk kili-kili, Fi kili- *armpit*.

PHN -lalak, RukTa, RukMn lalake *child*, Pai lalak *child* (term used by elders), lalak-an *little finger*, Puy lalak *young*, Tha ?a-^oä:ðak *child*, SblBt ?a-la:lak *offspring*, Png gi-la:lak *children, direct descendants*. Puy lalak (also) *children* (plural of alak *child* from PHN w₂aNak) seems to belong here, but if so is in a suppletive relation to its singular associate.

PHN pilay, RukBd ma-pilai, Pai ma-pilay, Tag, Ilk pi:lay, Ivt piday, Itb pilay *lame*.

PHN [tT]alam, Sed -talaŋ, Paz mi-talam *to run*, Sng t/um/alaŋ *to run away*.

To these it may eventually prove possible to add with assurance the following which involve what now appear to be inexplicable irregularities:

(?) PHN siliw, Paz siliw *running noose*, Tag si:loq, Ilk si:lu *loop, lasso*.

(?) PHN walíS, RukBd válisi *tooth*, Paz walis *tusk*, Saw ⁺li *canine teeth (horses, pigs)*, Kmb uli *tusk*.

In this connection Tsuchida (1976:143) offered the following comparison as implying a PHN [tT]alam, here reinterpreted:

PHN [tT]aNam, Kan ku-a-tanáme?, Sar m-aku-a-talame?, Tso oo-thom⁺, Bun tanam-un, Ami mi-tanam (A: n/l), Paz mu-talam, Sai šan-talam, Kvl talam (Dissimilation: l/n before m), Aty t/m/alam, Itb taxam (Dissimilation: x from l/n before m) *to taste*. It is attractive to treat this comparison as containing the element that appears doubled in Tag namnam *taste*, Tbt namnam *to taste with the lips*, the first part being a prefixed element *[tT]a- such as appears in Dempwolff's reconstruction *ta(n)kub *to cover* when considered in relation to his *kubkub *to cover*. It is possible that the two instances of dissimilation that this reconstruction requires themselves reflect a single dissimilation in a proto-language that formed a doublet with the reconstruction made here.

The following are the instances of correspondences in final position that can be assigned to *l:

PHN bakal, Puy vakal *a kind of knife*, Pai vakal *dagger*, Tag bakal *iron-tipped stick used in rice-cultivation*.

PHN bukul, RukMg ma-bkulu, Paz ti-bukul *hunchback*, Kvl buqul *knot*, Ilk bu:kul *swelling, protuberance, bump*, Mal bɔŋkòl *bump, hump*.

PHN bukel or pukel (with an appropriate analogical change), Kvl buqul (A: u/e), KlnKl pukel *knee*, KlnKy puk+l *bone*.

PHN buqel, Sed bql-it *leg*, WBM buqel *knee*, Seb buqul-buqul *ankle*.

PHN kawil, Kan m-ati-káini?, Sar m-ari-kaili?, RukMn -kaili *caught on thorns*, RukMg -kvili *caught by vines*, Bik, Hlg ka:wil *hang*, Seb kawil-kawil *hangnail*.

PHN ta?pil, Kan sia-tapíni? *patch*, Ami mi-tapil *to patch*, Bik taqpil *to patch*.

In regard to *l in final position Tsuchida (1976:143) cited Tag kawil *fish-hook* in connection with *kawil above and has found Dahl's suspicions confirmed by the discovery of Puy kawil-an *fishhook*. The Philippine words cited above seem to show a better semantic fit with the Formosan words than with the words for *fishhook*.

There are two instances that involve metathesis which confuses the issue as to whether a correspondence in final position is involved, though there appears no reason to doubt the cognation:

PHN [bp]eŋel or [bp]eleŋ, Sai pelen *deaf*, Hlg buŋul *deaf*, Mal biŋal *temporarily hard of hearing*. Whichever labial is original, the other is due to an analogical change.

PHN Zawil, Sar ma-sail-a?, RukTo ma?a-davili, Puy a-dawil, Btk, TbwK qa-lawid *far*. All non-Formosan words exhibit a metathesis (*Z-1/1-Z).

Initial correspondences of the same type as the medial and final correspondences assigned above to *l seem to be numerically adequate:

PHN laŋsi?, Puy laŋsi? *smell of burnt rice*, Ilk laŋsi *smell of certain fish, certain skin diseases, putrid blood, etc.*

PHN la(m+)lam, Sar ma-lalame?, Tso a-hmohmo *accustomed*, Png lamlam *become accustomed*.

PHN lekeC, Kan ma-ta-nekéce? *sticky*, Mal l+kak *adhere*.

PHN leklek, Puy -leklek, Itb leklek *to tickle*.

PHN lemek, Puy a-lmek *fine, soft*, Png an-lemek *become soft, soften*.

PAN libu?, limbu?, Paz libu? *hedge, fence*, AtySq libu? *circle, enclosure, trap*, Tag limbo *moon halo*, BM libu *surroundings*, Bar libu *ring around sun or moon*, mo-libu *sit in a circle*, Fi ma-levu (? e/i: inexplicable, but see below) *fish-weir enclosure*.

PHN libu?, Kan ni'ívu?, Sar livu?u?, RukBd líbu, RukMg libúu, RukTo libú, RukMn livu, Pai livu, Sai libu? *wild pig's grass-lined den*, Aty, Sed libu? *den, nest*, Biak niw *pig's lair*.

It is not at all unlikely that the last two comparisons belong to the same cognate set. They have been separated to provide ease in considering them either separately or together. The last comparison indicates that the regular reflex in eastern Indonesia may be *n*. Saw n+Bu *nest* looks as though it should also be included, but its initial consonant and first vowel offer serious difficulties; the Sawu word can more easily be associated with Numfor *niy+w pig's lair*. It is of course not unlikely that together they exemplify an uncontracted doublet of the last etymon above. It would however be most interesting if this uncontracted form might explain the aberrant vowel in Fi ma-levu.

There are two instances, both of which involve a PAN prefix *qali-, in which it is difficult to decide whether the *l should be regarded as initial or medial because the initial syllable is not reflected by many languages. The prefix appears in full in the following:

PHN qalibanḡaḡ, Puy Halivaḡḡaḡ, Seb qalibanḡaḡ *butterfly*.

The two instances are:

PHN (qa)lima[tT]ek, Kan ʔanimetéḡ-aʔ (A: e/a) *creek leech*, Sar ʔalimetek-aʔ (A: e/a) *paddy leech*, RukBd limáteke, RukTo limátake (A: a/e) *mountain leech*, RukMg lmátḡ, lmátḡe (Tsuchida) *leech*, RukMn limateke *small ground leech*, Pai limačḡek *mountain leech*, Puy limatek *paddy leech*, Tag lima:tik *leech*, Mer dimátika *small leech*.

PHN (qa)lim+Caq, Kan nimécaʔeʔ *paddy leech*, Ami la-lintaḡ *mountain leech*, Isg alimta *a kind of very large leech*, Mal (h)alintah, lintah *leech*, Bar alinta *leech*.

In the following there appears to be good reason to reconstruct a doublet, one with initial *l and the other with initial *n. Otherwise we must face the task of choosing between the about equally complicated possibilities of many independent instances of dissimilation and many instances of partial assimilation.

PHN luaḡ, Sar ta-i-luaḡe *female pygmy deer*, RukBd lúḡe *cow*, Blw, KlāG luwaḡ *carabao*.

PHN nuaḡ, RukMg nuḡḡe, RukTo nwáḡe, RukMn nuaḡe *cow*, Tha qnuwan *deer*, carabao, Paz nuaḡ *carabao*, cow, Ilk nuaḡ, Agta q+nwaḡ, Atta nua:ḡ, BonG, Ifg, Ibl nuwaḡ, Isg nuwa:ḡ, ItgB nuwaḡ, KnkN, KlnKy nuwaḡ, KlnKl newaḡ *carabao*.

Kan ʔi-núḡeʔ *female deer*, BunS ha-nvaḡ, BunNC qa-nvaḡ *deer*, carabao can be associated with either reconstruction, since both Kakanabu and Bunun merge *l and *n. Sai ha-nuan *horse* belongs here under a hypothesis that an assimilation (n/ḡ) occurred.

The evidence for a PHN luaḡ is somewhat strengthened by the appearance of support in a number of defunct Formosan languages: Fav loan *buffalo*, Bab loan, Pap loan, luang, Hoa loan, loang *cow*, carabao, Sir louang *ox*. Although it might be said that *nuaḡ is somewhat favoured to be the older form by its slightly greater distribution, it is difficult to contemplate the limited distribution of the two cognate sets and the fact that either one can be derived easily from the etymon of the other without regarding them as evidence for a Proto-Formosan on one hand and for a closer relationship between such a Proto-Formosan and Proto-Philippine as continuing a Proto-Hesperonesian.

There seems to be general agreement that the initial correspondence assigned here to *l is valid. However Dahl has argued that this initial correspondence

should be assigned - under his symbol *Da.L (Dahl 1976:74f and 1981:101ff) - to the same correspondence as the one that has been assigned here to *N. His grounds appear to be (1) that the instances exemplifying the medial and final correspondences assigned by Tsuchida to *l (under his then symbol *L) were not convincing; (2) that the instances exemplifying the initial correspondences like those for medial and final *N cited by Tsuchida could be explained as due to assimilation to a following nasal; and (3) that therefore the initial correspondence assigned to *l was in complementary distribution with the medial and final correspondence assigned to *N. Since we have introduced comparisons not cited in Tsuchida 1976, it is not clear how Dahl would view the problem now, particularly since some of the comparisons show non-Formosan cognates with laterals unassimilated to a following nasal.

Furthermore there is a small collection of comparisons with initial correspondences like the medial and final correspondences assigned to *N which do not lend themselves easily to being explained as resulting from instances of assimilation:

PHN Na[tT]aD, RukTa latáDe, RukBd látaDe, RukMn latade, Sai latar *outside*, Bik na:tad *front yard*.

PHN NaCeŋ, Kan natéŋe?, Sar lateŋe?, RukBd láceŋe, RukMg lcéŋə, RukTo lacéŋe *vegetables*, RukMn laceŋe *Solanum nigrum*, Pai laceŋ, Ami lateŋ, Ilk, Itb nateŋ *vegetables*, Ivt nateŋ *Solanum nigrum*, Bik natuŋ *taro, taro leaves*.

PAN Nuka?, Kan núka?, Tso h'o-h'o, RukMg ma-lku-lkáa, Bun nuka? *tumour*, Ami luka? *wound*, Paz luka? *scab*, Sed lu-qih, lu-qah, Itb nuka *wound*, BM nuka *skin eruption, scabies*, Mal luka (l/n, see below), TBT luha, ma-luha (l/n, see below) *lightly wounded, split apart, perforated*, Paul nua *itch, scabies*, Sam manu?a *wounded*, To manuka *be killed (of chief, sovereign)*.

With the first etymology immediately above might be associated the following words: Jav natar, latar, Snd latar *yard, land around the house*. However Bal natah *yard* suggests that Jav natar reflects *R₃₄. Snd latar could be explained from a *lataR₂, an etymon with initial *l correlative to *DataR₂ (Dempwolff 43) in the same way as *lemlem *dark* (Dempwolff 95) is correlative with *DeDem *dark* (Dempwolff 43). Jav latar might then be a Sundanese loanword. Although Snd latar would formally and semantically match the Rukai words - for Snd r is the outcome of *D as well as *R₂ - the Bikol word has the advantage of being unambiguous. The further possibility that there may have been an interplay between a *[lL]ataR₂ more or less homosemantic with a *NataD is perhaps suggested by the association of Mal, Mad natar, Jav latar *background, basic colour*. Finally one must also consider the possibility that Snd, Jav latar result from a back formation from Snd p̄-latar-an *land on which a building stands, front yard*, Jav p̄-latar-an (*large open (fore) square* which through a dissimilation (l-n/n-n) could be from the same *p̄-latar-an that is indicated by BalNoble p̄-latar-an beside natar *yard*. Under the hypothesis of a dissimilation and back formation Jav, Snd latar could be associated with *NataD. Jav natar likewise could be associated, indeed more directly, but Bal natah would remain problematic.

Since only Malay and Toba Batak exhibit initial l in the correspondence, it appears to be simplest to interpret them both as due to a dissimilation in just such a form as a *ma-Nuka? which might have been the source of Toba ma-luha with luha resulting by the analogical change often called back formation. Malay luka could have resulted in the same way, but, if so, at an earlier time, since there is no occurrent *ma-luka. In fact it is not unlikely that we might be dealing with a single instance of a dissimilation followed by a back formation that occurred in the last common proto-language of Malay and Toba.

Dahl (1981:105f) has suggested the words supporting the reconstruction of *NaCeŋ above could be associated with Dempwolff's *lateŋ *nettle* and offers the suggestion that the nettle is sometimes eaten as a vegetable. The English word *nettle* means a type of herb, and its young leaves are sometimes used in or as food, at least in Europe; on the other hand, all the cognate words of Dempwolff's *zalateŋ/lateŋ, so far as can be established, refer to a *nettle tree*, *Laportea spp.*, whose leaves have never been reported to be eaten. On the basis of the Formosan evidence we can now reconstruct *LaCeŋ, SaiTa käh-Laseŋ, SaiTu ra-aseŋ (from *[zZ]a-LaCeŋ), Puy L-iŋ-aTeŋ (M: ŋ-n/n-ŋ; A: n/l), Ami L-il-ateŋ *nettle tree*, *Laportea pterostigma*. The probable disconnection of the two etyma is favoured by the fact that Ivatan nateŋ above is found beside Ivt hateŋ (with h regularly for Dempwolff's l, here our *L) *a tree whose leaves on touch cause smarting pain and sores*, since the latter can hardly be dissociated from Dempwolff's *lateŋ. One might consider the possibility of a common source of the two Ivatan words through, let us say, a dialectal partial assimilation of an early Ivatan or pre-Ivatan initial lateral to the final nasal with subsequent semantic specialisation of one of the resultant doublet members, but this solution seems unnecessarily complicated.

The possibility of the dissimilation of a nasal as well as the partial assimilation of a lateral in relation to a following nasal can lead to uncertainty in the interpretation of the following comparison:

PHN [lN]awuŋ, Pai lauŋ, BunNCS navuŋ *shade*, Sai, Tbt lauŋ *shadow*, Mal nauŋ *shadowing, shelter*.

The purpose of this article has been to indicate that the claim that the reflexes assigned by Tsuchida to *N and *l (= TS. *L) can be regarded as being in complementary distribution faces rather strong opposing evidence. At the same time it should be noted that the area of nasals and laterals has begun to show complications that have not been dealt with here simply because the treatment would involve too many tangents. Naturally it remains possible that the collection of evidence presented here may ultimately receive another interpretation, but it hardly seems likely that the hypothesis of complementation will come to be justified. What seems clearly indicated is that careful investigation is called for to solidify our reconstructive hypotheses in this area.

In the course of the discussion we have presented reasons for believing that Proto-Austronesian had a distinction between a probably voiced lateral (*l) and a flap (*L), the latter articulation being found, thus far at any rate, only in the Formosan languages. The merger of *l and *N is found in all the Formosan languages with the apparent exception of Kavalan; the evidence for l from *l in this language is meagre, being limited to the two words (bilauŋ *to count*, talam *to taste*) cited above. Since there is little evidence that can be interpreted as favouring setting Kavalan up as a subgroup by itself, the merger of *l and *N can be regarded as evidence for a Proto-Formosan that had an isogloss separating merging and non-merging dialects, the latter continued only by Kavalan.

ADDENDUM

To the above evidence for the distinction between *l and *N the following can now be added:

*balluku?, AtyMx baluku?, Kvl bnuqu *winnowing basket*, Bontok balluku *small head-basket*. The Kvl word suggests at least a partial merger of *l with *n.

*lawlaw, Puy lawlaw, Tongan lolo *oil*.

ABBREVIATIONS OF LANGUAGE NAMES

AmF - Fataan Ami	Ngj - Ngaju(-Dayak)
Aty - Atayal	Pai - Paiwan
AtyMb - Mabatoan Atayal	PAN - Proto-Austronesian
AtyMx - Mayrinax Atayal	Pap - Papora
AtySq - Squliq Atayal	Paul - Paulohi
Bab - Babuza	Paz - Pazeh
Bal - Balinese	PHN - Proto-Hesperonesian
BalNoble - Noble Balinese (Dutch "voornaam")	Png - Pangasinan
Bar - Baree	Puy - Puyuma
Bik - Bikol	Ruk - Rukai
BM - Bolaang Mongondow	RukBd - Budai Rukai
Btk - Batak (Philippine)	RukMg - Maga Rukai
Bun - Bunun	RukMn - Mantaoran Rukai
BunN - Northern Bunun	RukTa - Tanan (Tainan, Dainan) Rukai
BunNC - Northern and Central Bunun	RukTo - Tona Rukai
BunS - Southern Bunun	Sai - Saisiat
Fav - Favorlang	SaiTa - Taai Saisiyat
Fi - Fiji	SaiTu - Tunggho Saisiyat
Hoa - Hoanya	Sam - Samoan
Hlg - Hiligaynon Bisayan	Sar - Saaroa
Ilk - Ilokano	Saw - Sawu
Isg - Isneg	SblBt - Botolan Sambal
Itb - Itbayat	Seb - Sebu
ItgB - Binongan Itneg	Sed - Sediq
Ivt - Ivatan	SedTn - Tongan Sediq
Jav - Javanese	Sir - Siraya
Kan - Kanakanabu	SL - Samar-Leyte Bisayan
KlaG - Guinaang Kalinga	Snd - Sundanese
KlnKl - Keleyqiq Kallahan	Sng - Sangirese
KlnKy - Kayapa Kallahan	Tag - Tagalog
Kmb - Kambara (Sumba)	TBt - Toba Batak
Kvl - Kavalan	Tha - Thao
Mad - Madurese	TbwK - Kalamian Tagbanwa
Mal - Malay	To - Tonga
Mer - Merina	Tso - Tsou
	WBM - Western Bukidnon Manobo

REFERENCES

DAHL, Otto Chr.

1976 *Proto-Austronesian*. 2nd edition (1st edn 1973). Lund: Studentlitteratur.

1981 *Early phonetic and phonemic changes in Austronesian*. Oslo: Universitetsforlaget.

DEMPWOLFF, Otto

1938 *Austronesisches Wörterverzeichnis*. *Zeitschrift für Eingeborenen-Sprachen*, 19tes Beiheft. Hamburg.

ECK, R. van

- 1876 *Eerste proeve van een Balineesch-Hollandsch woordenboek.* Utrecht: Kemink.

LI, Paul Jen-kuei

- 1977 *The internal relationships of Rukai.* *Bulletin of the Institute of History and Philology, Academia Sinica* 48/1:1-92.
- 1982 *Kavalan phonology: synchronic and diachronic.* In Rainer Carle et al, eds *Gava^c: studies in Austronesian languages and cultures dedicated to Hans Kähler*, 479-495. Berlin: Reimer.

NOTHOFER, Bernd

- 1975 *The reconstruction of Proto-Malayo-Javanic.* *Verhandelingen van het Koninklijk Instituut voor Taal-, Land-, en Volkenkunde* 73. The Hague: Nijhoff.

OGAWA, N. and E. ASAI

- 1935 *The myths and traditions of the Formosan native tribes.* Taihoku.

TSUCHIDA, Shigeru

- 1976 *Reconstruction of Proto-Tsouic phonology.* *Study of Languages & Cultures of Asia & Africa, Monograph Series No.5.* Tokyo.

LINGUISTIC VARIATIONS OF DIFFERENT AGE GROUPS IN SOME FORMOSAN LANGUAGES

Paul Jen-kuei Li

1. INTRODUCTION¹

Many linguists work with only one informant for each dialect they investigate. The shortcoming of such a methodology is that a heterogeneous speech community is misrepresented as a completely homogeneous speech community. By so doing we are apt to miss extremely valuable information for historical linguistics. Because language may vary according to the speakers' social characteristics and to the social context, we ought to consult speakers of both sexes, different social classes, various age levels, formal and informal speech, etc. in order to get a more realistic picture of the speech community.

While doing field work on the Formosan languages, I have observed from time to time that in some of these languages, speakers of different age levels may have different sound systems. Many interesting examples of sound changes in progress can be drawn from the variations of speech forms exhibited in the speech of different age groups in some of the dialects.

Age certainly plays an important role in language structure and change. The linguistic variations of different age levels seem to have significant implications for historical linguistics. The following questions can be raised:

(1) Do linguistic variations indicate that the language is in a state of change? In general, older speakers retain more archaic forms, whereas younger speakers tend to produce innovative forms. (2) Does the same group of languages or dialects have the same tendency and direction of change? (3) Is the overall direction of sound change towards simplification, as suggested by most historical linguists? (4) Is sound change phonetically abrupt, but lexically gradual, as first suggested by William Wang (1969)? (5) Is language change faster and thus more readily observable than the structuralists, such as Charles Hockett (1958: 384, 439-444; 1965), have assumed? It seems to me that all changes in progress can be observed through the variations of different age groups within the same dialect. (6) Can language change be explained as a process of language acquisition, as suggested by Noam Chomsky and Morris Halle (1962)? (7) Do all sounds change at the same rate of speed, or do different types of sounds change at different rates of speed? We need substantial empirical evidence from various languages to answer these questions positively or negatively. Formosan languages provide an excellent laboratory for such linguistic investigations.

I (Li 1982) have recently made an extensive study of the linguistic variations of different age groups in the Atayalic group. This study indicates that it may take up to 30 or 40 years to complete a change and that those speakers

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 33-50. *Pacific Linguistics*, C-94, 1986.

© Paul Jen-kuei Li

between the oldest whose speech shows no change and the youngest in whose speech the change has been completed are the ones whose speech best shows the process of change.

In this paper I shall try to show how various Formosan languages change through time, as manifested in the variations of different age groups. When considerable amounts of data become available, we shall be in a better position to answer the questions concerning language history raised above.

2. VARIATIONS OF DIFFERENT AGE GROUPS IN SOME FORMOSAN LANGUAGES

2.1 The variants *r* and *h* in Mantauran

In the Mantauran dialect of Rukai, /*r*/ is gradually changing to /*h*/ as exhibited in the speech of different age levels; see Table 1 below.

Roughly speaking, Mantauran speakers in their 60s generally retain /*r*/ in all or most of the lexical forms, whereas speakers under age 40 or so have substituted /*h*/ for /*r*/ in all or most of the forms. But this varies from speaker to speaker. While a 65-year-old man named takanao (Liáng Jīn-chiāng² 梁金鏞 in Chinese) and a 60-year-old woman named əɾəŋə (Liú Yüè-jī 呂玉枝 in Chinese) retain all the /*r*/s in their speech, a 61-year-old woman named dipolo (Lín Chūn-jiāng 林春江 in Chinese) retains only 52% of the total number of the occurrences of /*r*/ investigated. A 53-year-old man named lanpao (Guān Chīng-jiāng 關清江 in Chinese) and a 41-year-old man named takanao nəŋanə (Wú Wǔ-shūng 吳武松 in Chinese) retain no /*r*/s in their speech; all /*r*/s have changed to /*h*/s. Some speakers have free variation between /*r*/ and /*h*/ for some of the items. The data seem to indicate that female speakers are somewhat more conservative than male speakers. Some Mantauran women in their late 30s still retain /*r*/ in a few lexical forms, as shown in Table 1. Thus it is taking about 25 years to complete the change *r* > *h* in the whole speech community of Mantauran.

The two consonants, /*r*/ and /*h*/, are phonetically rather different. It seems a bit strange for /*r*/ to become /*h*/, however, I came across a 54-year-old man named arasə (Jīn Mào-rúng 金茂榮 in Chinese) who often produced a voiceless trill [r̥], which sounds like a phonetic transition [hr] between the voiced trill [r] and the voiceless glottal fricative [h]. Most older speakers produce the voiced trill [r].

The Mantauran speakers are aware of the sound change. The older speakers say that the "standard" pronunciation should be /*r*/ and that only "children" pronounce it as /*h*/. A 36-year-old woman named saidai (Liú Guèi-lán 呂桂蘭 in Chinese) who has both /*r*/ and /*h*/ in her speech said that the tone is more "emphatic" to pronounce /*r*/ instead of /*h*/.

The change *r* > *h* has been completed in another Rukai dialect, Labuan; see Li 1977:85. It is not clear when the change took place in the dialect.

A similar sound change has been observed in the [ʔafari] dialect of Thao, another Formosan language. Thao has two dialects, Barawbaw and [ʔafari]. Barawbaw /*r*/ has become [ʔafari] /*h*/ only before /i/, e.g. /ʔari:na/ > /ʔahi:na/ ears; it has become /y/ elsewhere, e.g. /fu:rad/ > /fu:yad/ moon, /sa:ran/ > /sa:yan/ road (see Li 1983).

Table 1: Variations of rʷh by different age levels in Mantauran³

	T.m65	d.	T.A.	d.	A.	S.	L.	M.	T.m41	
	ə.f60	f59	m52	f61	m54	f36	f54	f37	L.m53	
1. ravoko	r	r	r	rʷh	h	h	h	h	h	lime
2. masarəŋərə	r-r ⁴	r-r	r-r	r-r	h-h	h-h	h-h	h-h	h-h	quiet
3. əmarə	r	r	h	r	h	h	h	h	h	moon
4. tarjo	r	r	h	r	h	h	h	h	h	star
5. əkarələ	r	r	r	h	h	h	h	h	h	river
6. tarədima	r	r	rʷh ⁵	r	h	h	h	h	h	flood
7. əəəərə	r	r	h	r	h	h	h	h	h	thunder
8. sirəto	r	r	r	r	h	h	rʷh	h	h	flint
9. ʔaṣiroka	r	r	r	h	h	r	h	h	h	orange
10. tuvoro	r	r	r	r	h	h	h	h	h	bamboo shoots
11. kariəŋərə	r	r	r	r	h	h	rʷh	h	h	pigeon peas
12. ʔavarə	r	r	h	r	r	h	h	h	h	palm tree
13. tualiliʔor	r	r	h	r	r	h	h	h	h	rice gruel
14. valitoro	r	r	h	r	r	r	h	h	h	sticky rice
15. kiriti	r	r	r	h	r	h	h	h	h	burned rice
16. mayrəŋərə	r	r	r	h	h	r	h	rʷ(h) ⁶	h	sweet potato
17. siririʔi	r	r	r	h	h	r	h	h	h	stalk of taro
18. pirinji	r	r	r	r	r	h	h	h	h	branch
19. karana:ə	r	r	r	r	rʷr	r	r	h	h	cogan grass
20. tarəmunahə	r	r	h	r	h	h	h	h	h	pumpkin
21. tarəŋavaə	r	r	h	r	r	h	h	h	h	mushroom
22. ərətai	r	r	r	r	h	h	rʷh	h	h	poisonous mushroom
23. ciŋiriʔi	r	r	r	h	h	h	h	h	h	sprout
24. vokaərə	r	r	r	r	r	r	h	h	h	plant sp.
25. tavikara	r	r	h	r	r	r	h	h	h	plant sp.
26. karəmə	r	r	r	r	r	r	r	h	h	plant sp.
27. taraimu	r	r	r	h	r	r	r	h	h	plant sp.
28. tarəŋa:ə	r	r	(r)	r	r	r	h	h	h	green onion
29. ka:rara	r-r	r-r	r-r	h-h	r-r	h-h	{h-h r-r	h-h	h-h	pangolin
30. mavoroko	r	r	r	h	h	r	h	rʷh	h	monkey
31. ciarə	r	r	rʷh	r	h	rʷh	h	h	h	wing
32. ʔoraʔoro	r-r	r-r	r-r	r-r	r-r	h-h	h-h	h-h	h-h	frog
33. korəpaŋərə	r	r	r	r	rʷr	rʷh	h	h	h	toad
34. karəðolo	r	r	r	h	rʷr	h	h	h	h	egg
35. tarokoko	r	r	r	h	r	h	h	h	h	chicken
36. tikoray	r	r	r	h	rʷr	r	h	h	h	bamboo partridge
37. koroToŋo	r	r	r	r	rʷr	r	h	h	h	sparrow
38. tavorirakici	r-r	r-r	r-r	r-r	{r-r h-r	h-h	h-h	h-h	h-h	swallow
39. ʔaromamaraʔə	r-r	r-h	r-h	r-r	r-h	r-h	h-h	h-h	h-h	centipede
40. aəsərəpə	r	r	h	h	h	r	h	h	h	moth
41. aəsərələ	r	r	r	r	rʷh	h	r	h	h	tadpole
42. palikəroŋərə	r	r	r	r	h	h	h	h	h	waist
43. cokonoŋo	r	h	h	h	h	h	h	h	h	head
44. karəðərə	r-r	r-r	r-h	h-h	h-h	h-h	h-h	h-h	h-h	black spots on baby head
45. sarəmə	r	r	r	r	r	rʷh	r	- ⁷	h	lower jaw
46. toTərə	r	r	rʷh	r	h	r	h	h	h	throat
47. taəkərə	r	r	r	r	r	h	h	h	h	armpit
48. voro	r	r	r	r	r	h	h	h	h	small intestine
49. taTosərəkaə	r	r	r	r	h	h	h	h	h	large intestine
50. varəŋə	r	r	r	r	h	h	h	h	h	belly
51. o-si-varəŋə	r	r	h	h	h	h	h	h	h	pregnant

(cont'd over)

Table 1 (cont'd)

	T.m65 ə.f60	d. f59	T.A. m52	d. f61	A. m54	S. f36	L. f54	M. f37	T.m41 L.m53	
52. paracoŋəə	r	r	r	r	h	h	h	h	h	lung
53. varoŋovoŋəə	r	r	r	r	h	h	rʋh	h	h	internal organs
54. kavəcəɾəə	r	r	h	h	h	h	h	h	h	calf of leg
55. əɾəi	r	r	r	h	h	h	h	h	h	blood
56. ɾacələ	r	r	r	r	h	h	h	h	h	scar
57. kəɾimədəɾə	r-r		r-r	h-h	h-h				h-h	testicles
58. savərə	r	r	r	h	h	h	h	h	h	young man
59. marədəŋə	r	r	r	h	h	h	h	h	h	old man
60. malovərə	r	r	h	h	h	h	h	h	h	thin
61. ɾəŋətə	r	r	r	h	h	h	h	h	h	charcoal
62. o-ɾitihitisi	r	r	r	h	h	h	h	h	h	spark
63. varitəə	r	r	r	r	h	h	h	h	h	trousers
64. atərə	r	h	h	h	h	r	h	h	h	ring
65. romoko	r	r	h	h	h	h	h	h	h	fish-weir
66. vakərə	r	r	h	h	h	h	h	h	h	basket
67. sororo	r-r	r-r	r-r	r-r	h-h	h-h	{h-h r-r	h-h	h-h	slate bed
68. osopərə	r	r	h	r	h	h	h	h	h	shrill
69. o-siraovo	r	r	rʋh	r	h	h	h	h	h	dance
70. macaviri	r	r	h	r	h	r	h	h	h	fight
71. mavario	r	r	r	r	h	h	h	h	h	quarrel
72. (o)kərokoro	r-r	r-r	r-r	r-r	{r-r r-r	h-h	h-h	h-h	h-h	dig
73. marimoro	r-r	r	r	h-h	h-r	h	h-h	h-h	h-h	forget
74. odəŋərə	r	r	r	h	h	h	h	h	h	aim at
75. ototoro	r	r	r	h	h	h	h	h	h	drip
76. otoroŋo	r	r	r	h	h	h	h	h	h	peck
77. ʔosario	r	r	h	h	h	h	h	h	h	play
78. (o)ədodo	r	r	h	r	h	h	h	h	h	push
79. o-sərəpə	r	r	r	h	h	h	h	h	h	suck
80. (o)sisiri	r	r	r	h	h	h	h	h	h	tear
81. oərəcə	r	r	h	h	h	h	h	h	h	tie
82. mavisərə	r	r	h	h	h	h	h	h	h	break down
83. maraməə	r	rʋh	h	h	r	h	h	h	h	resemble
84. maridərə	r-r	r-r	r-r	h-h	h-h	h-h	h-h	h-h	h-h	fast
85. marilai	r	r	r	h	h	h	h	h	h	skinny
86. ʔadiriɾə	r	r	h	h	h	h	h	h	h	greedy
87. maravərə	r-r	r-r	r-r	h-h	h-h	h-h	h-h	h-h	h-h	happy
88. ʔinilavarəə	r	r	h	h	h	h	h	h	h	jealous
89. malavərə	r	r	h	h	h	r	h	h	h	admire
90. ʔaraka:vaʔiikai	r	r	r	h	h	h	h	h	h	late
91. maʔosaravə	r	r	h	r	h	r	h	h	h	scattered
92. masorərə	r-r	r-r	r-r	r-r	h-h	h-h	h-h	h-h	h-h	smooth
93. toramoro	r-r	r-r	r-r	r-r	r-h	h-h	h-h	h-h	h-h	serious
94. okəsərə	r	r	h	rʋh	h	h	h	h	h	sharp
95. ʔorokərə	r-r	r-r	r-r	r-r	r-r	{r-h h-h	{r-r h-h	h-h	h-h	loose
96. mavərəʔə	r	r	h	h	h	h	h	h	h	tired
97. a:na rə:a	r	r	h	h	h	h	h	h	h	there
98. iriə	r	r	r	r	h	h	h	h	h	left
99. tarəənəmə	r	r	r	r	h	h	h	h	h	six
No. of [r]:	114	106.5	78	59	31	18	8	1	0	
No. of [h]:	0	3.5	34	53	79	88	101	110	114	
% of change:	0%	3%	30%	47%	72%	83%	93%	99%	100%	

2.2 The variants γ > x in Skikun

As stated in my earlier report (Li 1982:180), the devoicing of the word-final velar fricative $[-\gamma]$ has been completed in both the Skikun and Mnawyan dialects of Atayal. The change $\gamma > x$ in word-initial and medial position is taking place in Skikun. However, data was unavailable to show how the devoicing process was taking place. I have collected more adequate data since then, as given in Table 2 below.

Apparently it will take many more years to complete the devoicing of $[\gamma]$ in the dialect. The 73-year-old woman named *yayuc koyuŋ*⁸ shows in her speech that changes have occurred in 15% of the cases collected, while the 29-year-old woman named *yaβuŋ ʔitay* shows only 51% of change. To complicate the matter, there is no clear correlation between the age and percentage of change. Some older speakers may show even more change than some younger speakers, for instance, the 69-year-old woman named *ʔipay naway* and the 61-year-old man named *pasəŋ ʔukis* show many more changes (72% and 76% respectively) than the 29-year-old woman and the 35-year-old man named *yukan tuliʔ* (51% and 66% respectively). Moreover, the 55-year-old woman named *yayuc nakaw* shows 91% of change. It is hard to predict the total number of years required to complete the change in the whole speech community of Skikun. The fluctuation between $[\gamma]$ and $[x]$ may be spread over a hundred years.

Table 2: The devoicing of initial and medial γ in Skikun

	YK f73	HY m49	YS m54	MS m56	P? m63	PS m64	?? f52	YT m37	SN m46	?N f69	PN m61	Y? f31	BW m37	YN f58	
Initial γ															
1. γon	γ	γ	γ	γ	γ	γ	x	γ	γ	x	γ	γ	γ	γ	<i>stream</i>
2. $\gamma\text{aya}^?$	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	x-x	γ - γ	γ - γ	x-x	x-x	x-x	<i>custom</i>
3. γamil	γ	γ	γ	γ	x	γ	x	$\gamma^{\sim}\text{x}$	γ	x	γ	γ	γ	x	<i>root</i>
4. γhap	γ	γ	$\gamma^{\sim}\text{x}$	γ	$\gamma^{\sim}\text{x}$	x	γ	x	x	x	x	γ	x	x	<i>seed</i>
5. γlux	γ	γ	$\gamma^{\sim}\text{x}$	γ	x	γ	x	x	x	x	x	x	x	x	<i>companion, together</i>
6. $\gamma\beta\text{ayan}$	x	x- γ	x- γ	x- γ	x- γ	x- γ	x- γ	x- γ	γ - γ x-x	x	x-x	x-x	x-x	x-x	<i>summer</i>
7. $\gamma\text{iqu}^?$	γ	γ	x	x	x	γ	x	x	x	x	x	x	x	x	<i>nit</i>
8. $\gamma\text{malu}^?$	x	x	x	γ	x	x	x	x	γ	x	x	x	γ	x	<i>sympathise</i>
9. γuquh	$\gamma^{\sim}\text{x}$	γ	γ	x	x	x	x	x	x	x	x	x	x	x	<i>banana</i>
10. $\gamma\text{ml}\gamma\text{ul}$	γ - γ	γ - γ x-x	x- γ	x-x	x- γ	γ - γ x-x	x-x	x-x	x-x	x-x	x-x	x-x	x-x	x-x	<i>grind</i>
11. γiyus	x	γ	x	x	$\gamma^{\sim}\text{x}$	x	x	x	x	x	x	x	γ	x	<i>intestines</i>
12. γlaquy	γ	γ	x	x	x	x	x	x	x	x	x	-	x	x	<i>female deer</i>
13. $\gamma\text{riy}\text{u}\text{ŋ}$	γ	x	γ	x	x	x	x	x	x	x	x	x	x	x	<i>back of neck</i>
14. $\gamma\text{lh}\text{i}\text{ŋ}$	γ	x	x	x	x	x	x	x	γ	x	x	x	x	x	<i>narrow</i>
15. $\gamma\beta\text{iy}\text{an}$	$\gamma^{\sim}\text{x}$	x	x	$\gamma^{\sim}\text{x}$	x	x	γ	x	x	x	x	x	x	x	<i>evening</i>
16. γlax	γ	γ	x	x	x	x	x	x	x	x	x	x	x	x	<i>right side</i>
17. $\gamma\text{ipu}^?$	x	$\gamma^{\sim}\text{x}$	x	γ	x	x	x	x	x	x	x	-	x	x	<i>female (animal)</i>
18. $\gamma\text{laqu}\text{ŋ}$	γ	x	x	x	x	x	x	x	x	x	x	x	x	x	<i>pheasant</i>
19. $\gamma\text{m}^?\text{ul}$	γ	x	x	x	x	x	x	x	x	x	x	x	x	x	<i>bury</i>
20. $\gamma\text{la}\beta\text{a}\text{ŋ}$	γ	x	x	x	x	x	x	x	x	x	x	x	x	x	<i>wide</i>
21. $\gamma\text{uru}^?$	x	x	x	γ	x	x	x	x	x	x	x	x	x	x	<i>duck</i>
22. $\gamma\text{mal}\text{i}\text{q}$	x	x	x	x	x	x	x	x	γ	x	x	x	x	x	<i>tear</i>
23. γqayus	$\gamma^{\sim}\text{x}$	x	x	x	x	x	x	x	x	x	x	x	x	x	<i>new</i>
24. $\gamma\text{suna}^?$	x	x	x	x	x	x	x	x	x	x	x	x	x	x	<i>breath</i>
25. γmlux	x	x	x	x	x	x	x	x	x	x	x	x	x	x	<i>ride</i>
26. γmatuk	x	x	x	x	x	x	x	x	x	x	x	x	x	x	<i>peck</i>
Medial γ															
27. $\text{sy}\gamma\text{ay}$	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	γ - γ	<i>separate, leave</i>
28. $\text{kin}\gamma\text{yan}$	γ	γ	γ	γ	γ	γ	γ	γ	γ	γ	γ	γ	-	x	<i>widow</i>
29. $\text{yay}\text{i}\text{h}$	γ	γ	γ	γ	γ	γ	γ	γ	x	γ	γ	x	-	γ	<i>baby girl</i>
30. $\text{my}\gamma\text{a}\text{l}$	γ	γ	γ	γ	γ	γ	γ	γ	γ	γ	x	γ	x	x	<i>painful</i>
31. $\text{cu}\gamma\text{an}$	γ	γ	γ	γ	γ	γ	γ	γ	γ	γ	γ	x	x	x	<i>tomorrow</i>

(cont'd over)

Table 2 (cont'd)

	YK f73	HY m49	YS m54	MS m56	P? m63	PS m64	?? f52	YT m37	SN m46	?N f69	PN m61	Y? f31	BW m37	YN f58	
32. (m)kaya?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	x	x	x	day after tomorrow
33. syalu?	Y	Y	Y	Y	Y	Y	Y	x	x	Y	Y	Y	x	Y ^{ux}	sympathise
34. rpayan	Y	Y	Y	Y	Y	Y	Y	Y ^{ux}	Y	x	Y	x	x	Y ^{ux}	whetstone
35. stnyan	Y	Y	Y	Y	Y	Y	Y	Y	Y	x	Y ^{ux}	x	Y ^{ux}	x	privy
36. tyuyal	Y	Y	Y	Y	Y	Y	Y	Y ^{ux}	x	Y	Y	x	x	x	three
37. mtkyil	Y	Y ^{ux}	Y	Y	Y	Y	Y	x	Y	Y	x	Y	x	x	dizzy
38. payay	Y	Y	Y	Y	Y	Y	Y	Y	x	x	Y	x	x	x	rice plant
39. ?imayal	Y	Y	Y	Y	Y	Y	Y	Y	x	Y	x	x	x	x	five
40. mayal	Y	Y	Y	Y	Y	Y	x	Y	x	Y	x	x	x	Y	take
41. kyis	Y	Y	Y	Y	Y	Y	Y	Y ^{ux}	x	x	x	x	Y	x	hemp plant
42. kmuyus	Y	Y	Y	Y	Y	Y	Y	x	Y ^{ux}	Y	x	x	x	x	shave
43. klyayan	Y	Y	Y	Y	Y	Y ^{ux}	Y	Y	Y	x	x	x	x	x	noon
44. puya?	Y	Y	Y	Y	Y	Y	Y	x	Y	x	x	x	x	x	navel
45. crayis	Y	Y	Y	x	Y	Y	Y	Y	x	x	Y	x	x	x	gaiter
46. myiyas	Y	Y	Y	Y	Y	Y	Y	Y	x	x	x	x	x	x	escape
47. ?ayiq	Y	Y	Y	Y	Y	Y	Y	Y ^{ux}	x	x	x	x	x	x	Miscanthus, thatch
48. ryyax	Y	Y	Y	Y	Y	Y	x	Y	x	x	x	x	x	x	mountain
49. wayi?	Y	Y	Y	Y	Y	Y	x	x	Y	x	x	x	x	x	sun
50. Bayah	Y	Y	Y	Y	Y	Y	x	x	Y	x	x	x	x	x	charcoal
51. pyiyas	Y	Y	Y	Y	Y	Y	Y	x	x	x	x	x	x	x	Escape!
52. tylas	Y	Y	Y	Y	Y ^{ux}	Y	Y	x	x	x	x	x	x	x	waterfall (small)
53. tyliq	Y	Y	x	Y	Y ^{ux}	Y	Y	x	x	x	x	Y	x	x	waterfall (large)
54. pyatuk	Y	Y	Y	Y	Y	Y ^{ux}	Y	x	Y	x	x	x	x	x	beak, woodpecker
55. byisa?	Y	Y	Y	-	Y ^{ux}	Y	Y	-	Y	x	x	x	-	x	reed of loom
56. qpuyu?	Y	Y	Y	Y	Y	x	x	x	x	x	Y	x	x	x	hairwhorl
57. byax	Y	Y	Y	Y	Y	Y	Y	Y	Y	x	x	x	x	x	testicles
58. layan	Y	Y	Y	Y	Y	x	x	Y	x	x	x	x	x	x	give up
59. tytap	Y	Y	x	x	Y	x	x	x	x	Y	Y ^{ux}	Y	x	x	fan
60. tyiyup	Y	Y	Y	Y	x	x	Y	x	x	x	x	x	x	x	sink
61. myop	Y	Y	Y	x	Y	Y	x	x	x	x	x	x	x	x	share one cup
62. ?uyiq	Y	Y	Y	Y	Y ^{ux}	Y ^{ux}	x	x	x	Y ^{ux}	x	x	x	x	sinew
63. cmuyu?	Y	Y	Y	x	Y	x	x	x	x	x	x	x	x	x	pound rice
64. tytux	Y	x	x	Y	Y ^{ux}	x	Y	x	x	x	x	x	x	x	earthquake
65. myiyaw	Y	x	x	x	x	Y	x	x	x	Y	x	x	x	x	lose
66. swayi?	Y	x	Y	x	x	x	x	x	x	x	x	x	x	x	kinship term
No. of [y]:	58½	48½	46	42½	41½	39	34	23½	20½	18½	16	11	7½	6	
No. of [x]:	10½	19½	24	25½	28½	31	36	45½	48½	50½	53	56	57½	64	
% of change:	15%	29%	34%	38%	41%	44%	51%	66%	70%	73%	77%	84%	89%	91%	

2.3 The variants bʋp and dʋt in Ishbukun

Bunun has two preglottalised voiced stops /b/ [ʔb] and /d/ [ʔd]. They are generally preserved in the word-initial and medial positions in all dialects. In the word- or syllable-final position, however, they are preserved only in the speech of older speakers in the Ishbukun dialect. Younger speakers tend to devoice them, as in Tables 3 and 4 below.

2.3.1 The variants bʋp

Table 3: Variations of bʋp (f or t) in Ishbukun⁹

	ʔ.f58 ʔ.m52	D. m48	ʔ. f38	ʔ. f27	S. f23	
1. kitɲab	b	b	p	b	b	<i>begin</i>
2. ɲabɲab	b	b-b	b-p	-	-	<i>cliff</i>
3. Xab	b	bʋp	p	-	-	<i>shinbone</i>
4. masubnuX	b	bʋp	p	p	b	<i>hate</i>
5. Xatub	b	bʋp	p	p	p	<i>trap</i>
6. si(su)suab	b	bʋp	b	b	f	<i>yawn</i>
7. laXaib	b	b	b	bʋp	p	<i>pass</i>
8. (ʔa)ʔabnum	b	p	p	t	p	<i>six (people)</i>
9. mabɛaX	b	p	-	p	p	<i>rainstorm</i>
10. minsunab	b	p	p	p	p	<i>burn</i>
11. misXaislab	b	p	p	-	p	<i>not burn up</i>
12. matukub	b	p	-	p	-	<i>cover</i>
13. maɲabɛan	b	b	b	p	p	<i>wide</i>
% of change:	0%	50%	66%	80%	80%	

The 13 lexical items listed in the table above are all the ones that I have found containing /b/ in word-final or syllable-final position in Ishbukun.

Speakers above age 50 retain /b/ in all lexical forms. In addition to the 58-year-old woman named ʔisut (Wáng Ā-pǐng 王阿丙 in Chinese) and the 52-year-old man named ʔadiman (Wǔ Wén-chuán 伍文傳 in Chinese), I have also worked with a 65-year-old woman named laɲus and a 74-year-old man named tulubus. They all retain /b/ in all the items. It should be noted that some younger speakers have not only devoiced /b/ to /p/, but have also changed it to /f/ or even to /t/. The occurrence of /t/ instead of /p/ in the form ʔatnum *six people* as given by the 27-year-old woman named ʔapin (Liǔ Ā-yuè 吕阿月 in Chinese) is due to assimilation to the following dental. Two younger speakers, toyuʔ (Sī Shíang-yún 司祥雲, male, aged 38) and sakuʔ (Wǔ Shìu-méi 伍秀美 in Chinese, female, aged 23) produced /f/ in the form sisusuab *to yawn*. Some younger speakers show the free variants bʋp in repetitions of the same lexical items.

I have found only one instance for the variants bʋʔ in the word-initial position in the Takituduh dialect. While the 56-year-old father named vatuʔ (Huáng Wàn-shēng 黄万生 in Chinese) produced the form bitvaq *thunder*, the 27-year-old daughter named kaut (Huáng Lì-chūn 黄麗春 in Chinese) produced ʔicvaq for the same item. All the four older speakers above age 30 I consulted gave initial /b-/, while the younger speakers under age 30 gave initial /ʔ-/. They all come from the same village, Takiqatuʔ.

2.3.2 The variants $d\check{v}t$

Parallel to the change $-b > -p$, there is also the change $-d > -t$, as exhibited in the speech of younger speakers of Ishbukun in Table 4 below.

Table 4: Variations of $d\check{v}t$ in Ishbukun¹⁰

	f58,65 m48,52	ʔ. f35	S. f30	ʔ. f27	S. f23	
1. Xud <u>u</u>	d	d	d	d	d	<i>drink</i>
2. matan <u>u</u> d	d	d	d	d	t	<i>prop up</i>
3. musl <u>u</u> d	d	t	d	t	d	<i>move over</i>
4. masuk <u>u</u> d	d	d	t	t	d	<i>close</i>
5. lisk <u>u</u> d	d		d	d	t	<i>stick</i>
6. XudX <u>u</u> d	d-d		d-d	d-d		<i>front neck</i>
7. masala <u>d</u>	d		t	d		<i>sharpen</i>
% of change:	0%	20%	25%	25%	40%	

2.4 The variants $t\check{v}c$ and $d\check{v}j$ in Toda

The changes from the dental stops /t, d/ to /c/ [tʃ] in the word-final position have been completed in all Sediq dialects; see Li 1981:254-255. The changes /t/ > /c/ and /d/ > /j/ [dʒ] are also taking place before /i/ or /y/ in some dialects, such as Toda and Toŋan, as manifested in the speech of younger speakers; see Li 1982:179. More adequate data have been collected for Toda, as shown in Tables 5 and 6 below.

Before /i/ or /y/, the change /d/ > /j/ applies to a few more lexical items than the change /t/ > /c/. In other words, the former change is going at a slightly higher rate of speed (from 37% to 11% of change) than the latter (from 16.7% to 9.5% of change). There is a clearer correlation between the age and percentage of change for the sound change /t/ > /c/ than that for the sound change /d/ > /j/. The difference in age is relatively small and not very obvious. Apparently the changes from dental stops to affricates are taking place very gradually and at this very slow rate over many decades; it may span more than a century.

Table 5: Palatalisation of /t/ before /i/ or /y/ in Toda¹¹

	R.B. f73	?.S. m72	W.S. m56	B.N. f47	B.T. m49	S.B. m25	
1. s <u>ə</u> t <u>i</u> ŋ	t	t	t	t	t	t	<i>hornet</i>
2. qu <u>t</u> <u>i</u> ʔ	t	t	t	t	t	t	<i>excrement</i>
3. <u>t</u> i <u>y</u> uʔ bawaʔ	t	t	t	t	t	t	<i>forefinger</i>
4. pu <u>t</u> iŋ	t	t	t	t	t	t	<i>knife</i>
5. to <u>t</u> iŋ	t	t	t	t	t	t	<i>steelyard</i>
6. qm <u>t</u> <u>i</u> ʔ	t	t	t	t	t	t	<i>defecate</i>
7. m <u>t</u> <u>i</u> ŋ	t	t	t	t	t	t	<i>fall, drop</i>
8. tm <u>t</u> <u>i</u> ŋ	t	t	t	t	t	t	<i>hit</i>
9. mru <u>t</u> iŋ	t	t	t	t	t	t	<i>muddy</i>
10. m <u>t</u> i <u>l</u> ux	t	t	t	t	t	t	<i>hot</i>
11. m <u>k</u> t <u>i</u> naʔ	t	t	t	t	t	t	<i>Bunun</i>
12. kr <u>ə</u> t-i	t	t	t	t	t	t	<i>Cut!</i>
13. sb <u>ə</u> t-i	t	t	t	t	t	t	<i>Beat!</i>
14. r <u>r</u> ə <u>t</u> -i	t	t	t	t	t	t	<i>Rub, Grind!</i>
15. t <u>q</u> <u>t</u> <u>i</u> ʔ	t	t	t	t	t	t	<i>fart</i>
16. <u>t</u> ya <u>q</u> ŋ	t	t	t	t ^{vc}	t	t	<i>crow</i>
17. <u>t</u> imuʔ	t	t	t	c	t	t	<i>salt</i>
18. <u>t</u> yad <u>ə</u> n	(t)	(t)	t	t	c	-	<i>dragonfly</i>
19. <u>t</u> i <u>y</u> ux	c	t	t	c	c	c	<i>small comb</i>
20. hu: <u>t</u> -i	t	c	c	t	c	c	<i>Suck!</i>
21. q <u>y</u> <u>t</u> -i	c	c	c	t	t ^{vc}	c	<i>Bite!</i>
No. of [t]:	19	19	19	18.5	17.5	16	
No. of [c]:	2	2	2	2.5	3.5	3	
% of change:	9.5%	9.5%	9.5%	11.9%	16.7%	14.3%	

Table 6: Palatalisation of /d/ before /i/ in Toda

	R.B. f73	S.B. m25	?S. m72	B.N. f47	B.T. m49	W.S. m56	
1. rḡədi?	d	d	d	d	d	d	<i>small fly</i>
2. dmīdil	d	d	d	d	d	d	<i>hold in hand</i>
3. tudīq	d	d	d	d	d	d	<i>drip</i>
4. chādīl	d	d	d	d	d	d	<i>heavy</i>
5. mdīras	d	d	d	d	d	d	<i>shout</i>
6. thādīl	d		d	d	d		<i>move dwelling</i>
7. hmādīl	d		d	d	d		<i>move things</i>
8. (d)ḡīma?	d	d	d	ɖvj	d	d	<i>bamboo</i>
9. hḡəd-i	d	d	d	d	d	ɖvj	<i>Cook (vegetables)!</i>
10. qhədīḡ	d	d	d	d	d	j	<i>hot spring</i>
11. tḡiyuḡ	d	d	j	d	d	d	<i>wasp</i>
12. qḡīdil	d	d	ɖvj	ɖv(j)	d	d	<i>woman</i>
13. qpḡḡi?	d	d	d	d	d	j	<i>earwax</i>
14. mtədīyal	d	d	d	d	j	d	<i>fight</i>
15. qəḡḡi?	j	d	d	d	d	d	<i>long</i>
16. tludīḡ	d		d	d	j	d	<i>middle finger</i>
17. səḡīq	d	d	ɖvj	d	d	j	<i>person</i>
18. rīḡi?	d	d	ɖvj	d	d	j	<i>handle (of knife)</i>
19. qḡīraw	d	d	d	j	ɖvj	ɖvj	<i>hawk</i>
20. kḡīyac	d	d	d	j	ɖvj	ɖvj	<i>cicada</i>
21. lḡud-i	d	d	j	d	d	j	<i>Connect!</i>
22. ḡīyax	d	d	d	j	j	ɖvj	<i>day</i>
23. ḡīyan	d	j	d	j	j	d	<i>daytime</i>
24. pḡīdiq	d	d	j	ɖvj	ɖvj	j	<i>Chinese cabbage</i>
25. buḡi?	d	j	j	ɖvj	d	j	<i>arrow</i>
26. ḡīraḡ	j	j	d	j	j	d	<i>dove</i>
27. kaḡi?	j	d	j	ɖvj	j	j	<i>fishnet</i>
No. of [d]:	24	21	20.5	19.5	19.5	15	
No. of [j]:	3	3	6.5	7.5	7.5	10	
% of change:	11.1%	11.1%	24.1%	27.8%	27.8%	37%	

I checked with two Toda speakers, *basaw temu?* (m49) and *bakan nawi?* (f47), in both February 1980 and August 1983, and they both showed slight changes in producing the dentals. *Basaw temu?* showed the change /t/ > /c/ in the lexical form *tiyux* > *ciyux comb* and the change /d/ > /j/ in the lexical forms *kadi* > *kaji fish net*, *diraḡ* > *jiraḡ dove*, *kḡījac* > *kḡīyac cicada*, *pḡīdiq* > *pḡīdiq ~ pḡījiq Chinese cabbage*. Surprisingly, *bakan nawi?* produced only *cyaquḡ* in 1980, yet she gave the free variants *tyaquḡ ~ cyaquḡ crow* in 1983, and she produced only the affricate /j/ in the lexical forms *tḡiyuḡ wasp*, *kaji* and *pḡījiq* in 1980, yet she gave only the dental /d/ or the free variants /d/ ~ /j/ in the forms *tḡiyuḡ*, *kadi ~ kaji* and *pḡīdiq ~ pḡījiq* in 1983. Her speech both in 1980 and in 1983 showed great fluctuation between the dentals and the affricates. It is possible, of course, that these variations did exist in *basaw temu?*'s speech in 1980 but did not turn up in the data I recorded.

2.5 The variants $r\ve$ and $e\emptyset$ in Tsou dialects

The main Tsou dialects are Tapangu, Tfuea and Duhtu (Tung 1964:2, Tsuchida 1976:83, Li 1978). The voiced retroflex continuant /r/ is retained only in the speech of the older speakers of Duhtu. It has become /e/, /y/ or zero, depending on the phonetic environments, in both Tapangu and Tfuea, as well as in the speech of younger speakers of Duhtu (Li 1978:281).

Tsuchida (1976:84-85) gave the following interesting sociolinguistic account concerning /r/ in Duhtu:

In the Duhtu village are found three types of speakers: (1) those who more or less constantly retain /r/; (2) those who do not have /r/ ...; (3) those who have free variants with and without /r/. Only a few of the speakers belong to the first type above and are mostly older than sixty years of age [in 1968]; they number probably less than ten. Although they pronounce /r/ consistently when asked to pronounce words in isolation, the /r/ sometimes does not appear when they tell a story The speakers who belong to type (2) are mostly women and those who are closely affiliated to Tfuea All the rest belong to type (3).

Later I found only one old man named avai tenayana (Jí Tàì-píng 吉太平 in Chinese, aged 74 in 1978), who still retained /r/ in his speech. It was difficult to consult him because of his poor health when I visited him again in 1982. Basically, then, /r/ has disappeared from Tsou by now.

There are some other speech variations exhibited by different age groups in the Tsou dialects. As it was stated in my earlier report (Li 1978:284):

Younger speakers of the Tapangu and Tfuea dialects tend to drop out the initial vowel /e/, which may have been historically derived from /r/ plus a vowel. For example, e'uho *leopard* in the older speakers' speech is 'uho in the younger speakers' speech.¹² In fact, the latter may have both variants.

Further examples for the variant $e\emptyset$ were listed in Li 1978:284. Unfortunately I did not consult speakers at various age levels then. This requires further investigation.

2.6 The variants -pʷ-k and -mʷ-ŋ in Skikun

Table 7: Variations of finals by different age groups in Skikun

(1)	SP f80	YK f71	BM m65	Pʷ m63	MW m50	PS m64	PN m61	MS m56	YS m54	BW m37	SN m46	Yʷ f31	ʷʷ f52	YN f58	YT m37	HY m49	WB m35	
1. qciyap	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	<i>opposite shore</i>
2. ʔiyup	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	<i>Chinese goshawk</i>
3. qatap	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	<i>scissors</i>
4. tytap	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	pʷk	-p	-p	-k	<i>to fan</i>
5. ɣhap	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	-k	<i>seed</i>
6. pshup	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	-k	<i>suck</i>
7. qurip	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	-p	-p	-p	-k	-k	<i>ginger pepper</i>
8. tyiyup	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	-p	-p	-k	-k	<i>sink</i>
9. hmap	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	pʷk	-p	-p	-p	-k	-k	-k	<i>stab</i>
10. qmuyup	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	-k	-p	-k	-k	<i>fold</i>
11. talap	-p	-p	-p	-p	-p	-p	-p	-k	-p	-p	-p	-p	pʷk	-k	-k	-p	-k	<i>eaves</i>
12. hmop	-p	-p	-p	-p	-p	-p	-p	-k	-p	-p	-p	-p	-k	-p	-k	-k	-k	<i>do magic</i>
13. qmalup	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	-k	-k	-k	-k	<i>hunt</i>
14. myop	-p	-p	-p	-p	-p	-p	pʷk	-p	-p	-p	-p	pʷk	-p	-k	-k	-k	-k	<i>share one cup</i>
15. msuyap	-p	-p	-p	-p	-p	pʷk	-p	-p	-k	-p	-p	-p	-p	-k	-k	-k	-k	<i>yawn</i>
16. kmiyap	-p	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	-k	-p	-k	-k	-k	-k	<i>catch</i>
17. mnep	-p	-p	-p	-p	-p	-p	-p	-p	-k	-k	-p	pʷk	-p	-k	-k	-k	-k	<i>to fish</i>
18. miyup	-p	-p	-p	-p	-p	-p	-p	-p	-p	-k	-k	-k	-k	-k	-k	-k	-k	<i>enter</i>
% of change:	0%	0%	0%	0%	0%	3%	3%	11%	11%	11%	14%	22%	31%	47%	50%	72%	100%	

(cont'd over)

Table 7 (cont'd)

(2)	SP f80	Pʔ m63	BM m65	SN m46	YK f73	YS m54	PS m64	BW m37	MS m56	HY m49	Yʔ f31	MW m50	PN m61	YN f58	ʔʔ f52	YT m37	WB m35	
1. (q)qom	-m	-m	-m	-m	-m	-m	-m	-m	-m	-m	-m	-m	-m	-m	-ŋ	-ŋ	-ŋ	<i>anteater</i>
2. syam	-m	-m	-m	-m	mʔŋ	-m	-m	-m	-m	-m	-ŋ	-m	-m	-ŋ	-ŋ	-ŋ	-ŋ	<i>pork</i>
3. qmtam	-m	-m	-m	-m	-m	-m	mʔŋ	-m	-m	-m	-ŋ	mʔŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	<i>swallow</i>
4. rom	-m	-m	-m	-m	-m	-m	-m	-m	-m	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	<i>needle</i>
5. qinam	-m	-m	-m	-m	-m	-m	-m	-m	-m	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	<i>peach</i>
6. hmham	-m	-m	-m	-m	-ŋ	-m	mʔŋ	-m	-ŋ	-ŋ	mʔŋ	-ŋ	-ŋ	-m	-ŋ	-ŋ	-ŋ	<i>grope</i>
7. { mtlom lmom	{ -m -m	{ -m -m	{ -m -m	{ -m -m	{ -m -m	{ -m -ŋ	{ -m -ŋ	{ -m -ŋ	{ -ŋ -ŋ	{ -ŋ -ŋ	{ mʔŋ -m	{ -ŋ -ŋ	{ -ŋ -ŋ	{ -ŋ -ŋ	{ -ŋ -ŋ	{ -ŋ -ŋ	{ -ŋ -ŋ	<i>burn</i>
8. yuhum	-m	mʔŋ	-m	-m	-ŋ	-m	-ŋ	-m	-m	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	<i>gall</i>
9. cmom	-m	-m	-m	-m	-m	-ŋ	-ŋ	-m	-ŋ	-ŋ	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	<i>wipe</i>
10. mktlium	-m	-m	-m	-m	-m	-ŋ	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	<i>run</i>
11. prahum	-m	-m	-m	-ŋ	-m	-ŋ	mʔŋ	-ŋ	-m	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	<i>lips</i>
12. tmalam	-m	-m	-ŋ	mʔŋ	-m	-m	mʔŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	<i>taste</i>
13. mnkum	-m	-m	-m	-ŋ	-ŋ	-ŋ	mʔŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	-ŋ	<i>dark</i>
% of change:	0%	4%	7%	18%	25%	32%	39%	39%	50%	57%	71%	82%	86%	86%	100%	100%	100%	

3. DISCUSSION OF THE SOUND CHANGES

3.1 Different types of sound changes progress at different rates of speed

It is obvious that different types of sounds change at different rates of speed. Some sound changes, such as /r/ > /h/ in Mantaoran, can be completed in less than 30 years, while other sound changes, such as /ɣ/ > /x/ in Skikun, may span a century or more.

A distinction should be made between "change in apparent time" (distributions across age levels) and "change in real time" (comparison of records for different periods) (Labov 1980:253). In this paper I generally refer to change in apparent time. Since there are no earlier reports for many of the dialects I have studied, it is impossible to compare different periods of time for these dialects today. Since the same adult speakers show slight changes in their speech as they get older, it can be inferred that change in apparent time is somewhat shorter than change in real time.

A phonemic merger, either complete or partial, takes about 30 years in apparent time. For instance, /r/ has completely merged with /h/ in Mantaoran in about 25 years, as seen in data of people roughly between the ages of 60 and 35; see Table 1 in section 2.1. The labial stop /-p/ has partially merged with the velar stop /-k/ only word finally in the Skikun dialect of Atayal in about 30 years (roughly between the ages of 62 and 32), and the labial nasal /-m/ has partially merged with the velar nasal /-ŋ/ word finally in about 40 years (roughly between the ages of 76 and 36); see Li 1982:169 and Table 7 in section 2.6. The voiced velar stop /g/ has partially merged with /-w/ after the vowels /a/ or /u/ in the Inago dialect of Sediq in about 30 years (between the ages of 69 and 38), and with /-y/ after the vowel /i/ in about 20 years (between the ages of 60 and 38); see Li 1982:174-176.

It takes a good number of years to complete the devoicing process of a certain consonant. For example, the devoicing of [ɣ] to [x] in Skikun is likely to span more than a hundred years (see section 2.2), and so is the devoicing of /-b, -d/ in Ishbukun (see section 2.3). The on-going change /-b/ > /-p/ was reported for the Mabatuʔan dialect of Atayal, in which no change was observed in the speech of speakers above age 70, and there is only less than 20% change in the speech of speakers in their 40s; see Li 1982:181.

Palatalisation also takes a long time. For example, /t/ > /c/ and /d/ > /j/ before /i/ or /y/ in Toda will be spread over a century; see section 2.4.

3.2 Factors responsible for sound change

Different factors may come into play in a certain sound change. Age is clearly the main determining factor for the following sound changes: /-p/ > /-k/, /-m/ > /-ŋ/ in Skikun, /-g/ > /-w, -y/ in Inago, /r/ > /h/ in Mantaoran, /r/ > /e, y/ or Ø in Duhtu. Sex is a minor factor in determining the speed of each of these sound changes.

However, different age levels do not indicate the direction and speed as clearly in the other types of sound change, such as the devoicing of the fricative /ɣ/ > /x/ in Skikun, that of the stops /-b/ > /-p/ in Mabatuʔan, /-b/ > /-p/ and /-d/ > /-t/ in Ishbukun, and the palatalisation /t/ > /c/, /d/ > /j/ in Toda. What factors, other than age and sex, are responsible for these sound changes? Could it be different styles of speech (such as formal and informal

speech, careful and casual speech), social classes, social context, or some other factors? Further study would be required before we could answer this question.

The style of speech is certainly an important factor in determining which variant will occur, at least in certain pairs of variants in Formosan languages. For example, /r/ is retained when the older Duhtu speakers pronounce words in isolation, but it sometimes disappears when they tell a story (Tsuchida 1976:84). Interestingly enough, the initial consonant /g-/ immediately preceding another consonant is generally lost when elicited in isolated words in the Squliq dialect of Atayal, but it may turn up in the Squliq texts that I have collected from the same informant, e.g. gmalu? ~ malu? *to sympathise*, gmluw ~ mluw *to follow, to ride* (Li 1980:365). A Mantauran speaker said that the tone is more "emphatic" if words are pronounced with /r/ instead of /h/ (see section 2.1).

The data as given in this paper and in Li 1982 are based on the pronunciation of words in isolation. Most speakers were not told and were thus unaware of the kinds of things I was actually observing.

NOTES

1. This paper was written with the support of the National Science Council Grant NSC73-0301-H001-04, Republic of China and further with a research grant from Harvard-Yenching Institute.
2. The Chinese Romanisation was originally designed by Yuen Ren Chao and a few other members on the Committee for a Study of Chinese Romanisation, adopted and publicised by the Ministry of Education in 1928. However, it was not popularly accepted by the public mainly because of the complicated rules to indicate different tones. Hence, the Ministry of Education set up a new committee to revise the spelling in February 1984. The author was on the Committee. The revised spelling system is adopted in this paper.
3. In the following tables, the initials stand for the native names of the informants, followed by their sex and age.
4. When a certain sound that undergoes a sound change appears twice in a lexical form, it is separated by the hyphen '- '.
5. The tilde '~' indicates free variation in the speech of the speaker. No attempt was made to record the frequency of each variant. Only the variant with extremely low frequency is put in parentheses.
6. See note 5 above.
7. When a certain item is not in the informant's vocabulary, it is indicated by '-'. When information is unavailable, it is left blank.
8. Most of the informants' names appeared in Li 1982:168,169. A few informants are added in Tables 2 and 7:

pasaŋ ʔukis (Lǐ Yǔng-shiúng 李永雄, m63);
 silan nokan (Chén Wén-jāng 陳文角, m46);
 (Jāng Tzan-chéng 張贊成, m56);
 ʔipay ʔagin (Lǐ Wén-yīng 李文英, f52);
 Yabuŋ ʔitay (Yè Lì-huā 葉麗花, f31);
 pasaŋ nabuʔ (Fāng Shīn-fā 方新發, m61).

9. The Ishbukun informants' names are:

- ?isut (Wáng Ā-bǐng 王阿丙, f58);
 ?aɛiman (Wú Wén-chuán 伍文傳, m52);
 ?akuan (Wú Ā-shīn 伍阿新, f35);
 daXu? (Wáng Lián-shèng 王連勝, m48);
 ?akuan (Shǐ Ā-fēng 史阿豐, f38);
 toyu? (SĪ Shiáng-yún 司祥雲, m38);
 saluŋ (Wáng Ā-gān 王阿甘, f30);
 ?apin (Liú Ā-yuè 呂阿月, f27);
 saku? (Wú Shiòu-mèi 伍秀美, f23).

10. All the informants listed in Table 4 are the same as the ones listed in Table 3 except Laŋus (Liú Ā-jū 呂阿珠 in Chinese, f65).

11. The informants' names in Tables 5 and 6 are:

- rabaw bawxi? (Jāng Wǔ-mèi 張五妹, f73);
 ?awi? suyan (Jāng Tiān-tsiuán 張珠泉, m72);
 walis suyan (Jāng Tiān-tsái 張珠財, m56);
 bakan nawi? (Hú Ài-mèi 胡愛妹, f47);
 basaw temu? (Jāng Wén-wàng 張文旺, m49);
 suyan basaw (Jāng Shùn-fú 張順福, m25).

12. The Proto-Formosan (PFN) form for the item is *likuɫaw *leopard* as attested in Atayal akl-i?, Sediq rkel-ic, Tsou (Duhtu dialect) r?uho, Kanavu ukúnau, Saaroa lukuɫu, Rukai, Paiwan, Puyuma Likulaw, Thao rukdaw, Saisiat Loklaw, Kavalan Luqəwaw, Ami Lukdaw. Hence the older speakers' speech retains the more archaic form in Tsou.

*Author and editors thank Han Kartawinata for his calligraphy!

REFERENCES

- BLOOMFIELD, Leonard
 1933 *Language*. New York: Holt, Rinehart and Winston.
- CHEN, Matthew Y. and William S.-Y. WANG
 1975 Sound change: actuation and implementation. *Language* 51/2:255-281.
- HALLE, Morris
 1962 Phonology in generative grammar. *Word* 18:54-72. Reprinted in Jerry A. Fodor and Jerrold J. Katz, eds *The structure of language*, 1964:334-352. Englewood Cliffs, N.J.: Prentice-Hall.
- HOCKETT, Charles
 1958 *A course in modern linguistics*. New York: Macmillan.
 1965 Sound change. *Language* 41:185-205.
- LABOV, William
 1972 *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

LABOV, William, ed.

- 1980 *Locating language in time and space*. New York: Academic Press.

LI, Paul Jen-kuei

- 1977 The internal relationships of Rukai. *Bulletin of the Institute of History and Philology (BIHP)* 48/1:1-92. Academia Sinica, Taipei.
- 1978 Variations in the Tsou dialects. *BIHP* 50/2:273-300.
- 1980 The phonological rules of Atayal dialects. *BIHP* 51/2:349-405.
- 1981 Reconstruction of proto-Atayalic phonology. *BIHP* 52/2:235-301.
- 1982 Linguistic variations of different age groups in the Atayalic dialects. In Kun Chang, ed. *Studies in linguistics presented to Dr. Fang-kuei Li on his eightieth birthday*. *Tsing Hua Journal of Chinese Studies* 14:167-191.
- 1983 Notes on Thao dialects. *Bulletin of the Department of Archaeology and Anthropology* 43:48-50.

SAPIR, Edward

- 1921 *Language*. New York: Harcourt, Brace & World.

TSUCHIDA, Shigeru

- 1976 *Reconstruction of Proto-Tsouic phonology*. Study of Languages & Cultures of Asia & Africa monograph series No.5. Tokyo.

TUNG, T'ung-ho

- 1964 *A descriptive study of the Tsou language, Formosa*. Taipei: Institute of History and Philology, Academia Sinica Special Publications, No.48.

WANG, William S.-Y.

- 1969 Competing changes as a cause of residue. *Language* 45/1:9-25.

LEXICAL DIFFUSION IN SANGIR

J.N. Sneddon

1. INTRODUCTION

In the Sangir language of northern Indonesia all non-nasal consonants were removed from word-final position through replacement by glottal stop or addition of the syllable əʔ.¹

Although the two processes were largely in complementary distribution in terms of the phonological environments in which they operated, this was not always the case and apparent irregularities present a difficulty for the traditional claim that sound change operates uniformly without exception. For instance, the syllable əʔ was generally added to words ending in *s, e.g. Proto-Austronesian (PAN) *huRas became Sangir uhasəʔ *wash*. But in some words *s was replaced by ʔ, e.g. PAN *nipis became Sangir nipiʔ *thin*. No phonological or other environmental factors can account for the different changes undergone by the two words above and, as is shown below, appeal to borrowing to account for such apparent irregularity, in this and many other instances, can be rejected.

This paper aims to show that (a) the two sound changes were competing methods by which final consonants were dealt with in Sangir, (b) both innovations spread gradually through the lexicon and (c) one innovation appeared first but was blocked from reaching the entire eligible lexicon by the other rule, this resulting in the apparent irregularity in the language today. The evidence presented also offers an explanation for the occurrence of doublets in Sangir, such as saʔuʔ *river* and saʔuhəʔ *gutter, riverbed*, which both reflect PAN *saluR, and dialect variation, such as Taruna dialect lewoʔ and Manganitu dialect lewohəʔ *young coconut*.

In this discussion evidence from some other languages needs to be considered. It is shown that not only were the changes the result of lexical diffusion in Sangir but that areal diffusion of the changes occurred also between Sangir and nearby languages which are also closely related to Sangir.

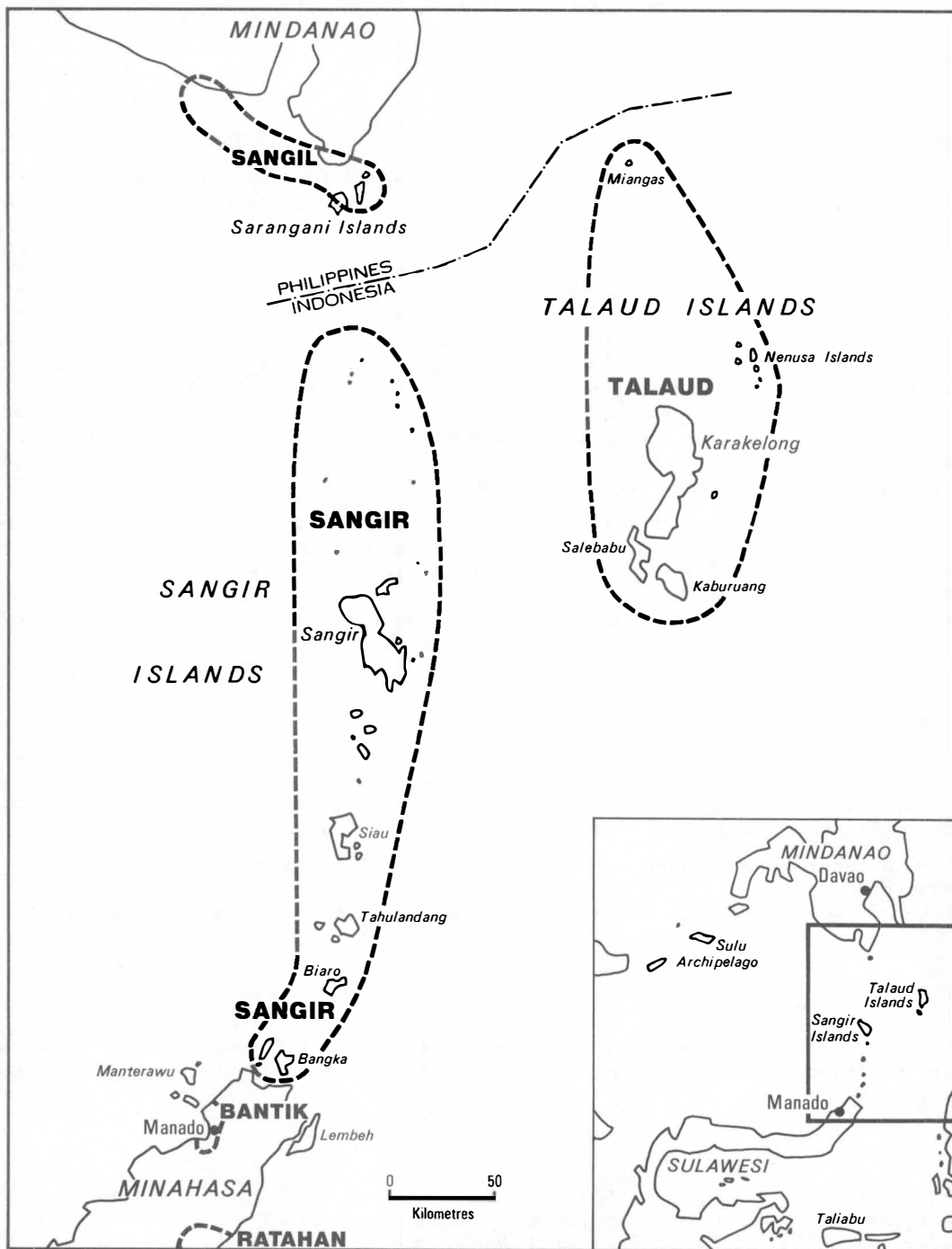
In general the languages most closely related to Sangir have been poorly documented to date. The paper therefore begins with some brief notes on the languages, before considering the historical changes which they have undergone.

2. THE SANGIRIC LANGUAGES

Sangir (San) belongs to a small group of Western Malayo-Polynesian languages, here called the Sangiric group, which probably link with the Philippine languages.²

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 51-71. *Pacific Linguistics*, C-94, 1986.

© J.N. Sneddon



Map: The Sangiric languages and their localities

A study of the interrelationships of the Sangiric languages has recently been carried out, which includes a statement of their phonological histories and a reconstruction of their parent language, Proto-Sangiric (Sneddon 1984). Proto-Sangiric (PSan) forms cited here are from the list of lexical reconstructions given in that work.

The Sangiric languages (see map, p. 52) are:

(a) San, spoken by about 180,000 people in the Sangir (Sangihe) Archipelago, which stretches from north of the eastern tip of the North Sulawesi peninsula towards Mindanao. Unless otherwise stated, San forms in this study are from the Manganitu (Mang) dialect.

(b) Sangil (Snl), spoken by about 10,000 people in coastal areas in the Cotabato and Davao provinces of Mindanao and also in the lower Sarangani peninsula and on the nearby Sarangani islands. In this study Snl forms are from the Sarangani dialect.³

(c) Talaud (Tal), spoken in the Talaud Archipelago, north-east of the Sangir Archipelago, by about 30,000 people. Tal forms given here are from the Salibabu dialect.

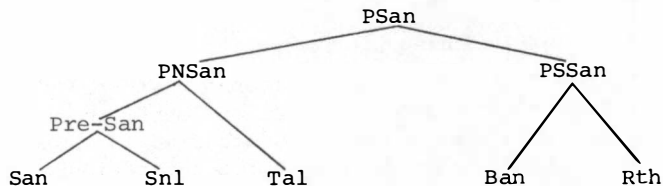
(d) Bantik (Ban), spoken in north-west Minahasa, the easternmost region of North Sulawesi, in about ten villages surrounding Manado, the provincial capital, and in several isolated villages further south.

(e) Ratahan (Rth), spoken by about 20,000 people in south-east Minahasa. The name Benteenan was sometimes used in earlier literature for this language.

Of these languages, San is by far the best known to linguists, publications including Adriani's grammar (1893) and Steller and Aebersold's dictionary (1959). San has also been used in wider comparative studies, e.g. Reid (1971), Charles (1974), Mills (1981).

There is sufficient high-quality evidence (which is presented in Sneddon 1984) to establish that Ban and Rth form one branch of the Sangiric group, the South Sangiric subgroup, continuing Proto-South-Sangiric (PSSan), while San, Snl and Tal form another, the North Sangiric subgroup, continuing Proto-North-Sangiric (PNSan). Within the North Sangiric subgroup San and Snl link closely together and could be regarded as two dialect groups, each with its own subdialects, of a single language.⁴ Because of their close relationship and the frequent need to refer to them together the abbreviation San/Snl is often used below. The historical changes discussed in this paper occurred largely within the period of their shared history, a stage here called Pre-Sangir (Pre-San), following Maryott (1978).

The genetic relationships of the Sangiric languages can be represented as follows:



Rth is geographically isolated from the other Sangiric languages and did not share in any of the innovations which spread through them. It is probable that the Rth community moved to their present location soon after their split with Ban.

The other Sangiric languages appear to have maintained close contact and a number of shared phonological characteristics are the result of areal spread.⁵ San and Tal are spoken on scattered island chains by predominantly fishing communities skilled in boating and there has always been a considerable amount of inter-island contact. San is spoken in at least one port in the Talaud islands (Team Fakultas Sastra Unsrat 1976-77). The Sangil people migrated to the Philippines from the Sangir Archipelago several hundred years ago (Maryott 1978) at which time they must have constituted simply another San dialect group. There has clearly been on-going contact between Ban and southern San communities and Ban shares a number of lexical items with the southern dialects Siau and Tahulandang (Thl) which are not found in the more northerly San dialects, such as Mang, Taruna and Tabukang, or in Snl. Also, there have long been Sangirese settlements on the Minahasan coast near the Ban speech area.⁶

3. TERMINOLOGY AND ABBREVIATIONS

Replacement of final consonants by ʔ is henceforth referred to as final consonant reduction, for which the abbreviation FCR is used. Any word-final consonant so affected is said to have reduced to ʔ . Thus the term 'reduction' for the purposes of this paper is clearly defined and distinguishes replacement of a final consonant by ʔ from other replacement or loss.

The additional or paragogic syllable, ə^{p} in San but phonologically different in some other Sangiric languages, has also developed independently in a number of other Sulawesi languages. It occurs in Makassarese, where Mills has called it 'the echo-vowel + [q] sequence' (1975:74). Adriani refers to it as an 'unaccented final syllable' ('*toonlooze eindlettergreep*') for San (1893:37). Such terms are too cumbersome for continual use; for ease of description and because of its precise definition the term paragoge will henceforth be used. This follows Maryott (1977) who uses the term to refer to the phenomenon in Snl. Pei (1966) defines 'paragoge' as: "The addition of a sound, letter or syllable to the end of a word, without etymological justification ... and without change of meaning in the word." There appears to be no derivative of the term 'paragoge' which refers to the paragogic syllable itself. Because of the constant requirement in this paper to refer to the syllable itself the term 'paragoge' is here used to refer to this paragogic or additional syllable and not to the process of its formation. The process of paragoge, as defined by Pei, will be referred to as paragoge addition, abbreviated to PA.

4. THE DIACHRONIC CHANGES

4.1 Final consonant reduction

FCR occurred in all Sangiric languages except Tal.⁷

In Rth $*t$ regularly reduced to ʔ , e.g. PSan $*\text{əpat}$ > Rth pa^{p} *four*, PSan $*\text{laŋit}$ > Rth laŋe^{p} *sky*, PSan $*\text{ikit}$ > Rth iki^{p} *tie*. No other consonants underwent reduction.⁸ It is possible that t-reduction occurred in PSSan, before its break-up into Ban and Rth. If so, then following the separation of Rth and its isolation from the other languages the change did not spread to other classes of sounds. Alternatively, t-reduction may have been an independent, parallel development in Rth after its separation from the other Sangiric languages. (Final consonant loss or reduction occurred separately in languages of various subgroups throughout Sulawesi.)

In Ban the voiceless stops *p, *t and *k reduced to ʔ: PSan *atup > Ban atuʔ *roof*, PSan *tiap > Ban tiaʔ *count*, PSan *takut > Ban takuʔ *afraid*, PSan *Ramut > Ban hamuʔ *root*, PSan *utak > Ban utaʔ *hair*, PSan *baluk > Ban baluʔ *sell*.⁹ No other consonants reduced to ʔ in Ban.¹⁰

Where FCR occurred in Rth and Ban the original consonant is not recovered before a suffix: Rth luwaʔ, Ban labaʔ *to cross (river)* (PSan *ləbat) → Rth luwaʔen, Ban labaʔen *be crossed*.

PSan voiceless stops reduced to ʔ in Pre-San, as reflected in San and Snl: PSan *atup > San, Snl atuʔ *roof*; PSan *ontap > San, Snl ontaʔ *bellows*; PSan *sepet > San, Snl sepeʔ *carry under arm*; PSan *ikit > San, Snl ikiʔ *tie*; PSan *utak > San, Snl utaʔ *hair*; PSan *manuk > San, Snl manuʔ *fowl*. There are some lexical items in modern San and Snl which have preserved p, t and k through paragoge addition; these are probably all borrowings (see Section 6 for discussion of these forms).

In Pre-San FCR also affected all other consonants other than nasals, i.e. voiced stops, *s, *l and *R (reflected as h in San and r in Snl), but only in a limited number of lexical items (see detailed discussion, with examples, in Section 6).

In San and Snl word-final glottal stop, whether reflecting PSan *ʔ or reduction of some other consonant, is replaced by another consonant preceding a suffix; by t if the preceding consonant is velar and by k elsewhere: San, Snl tiuʔ *blow* (PSan *tiup) → tiukaŋ *be blown*; San, Snl takaʔ *cover* (PSan *takap) → takaten *be covered*; San boheʔ *write* (PSan *boRet) → bohekaŋ *be written*, Snl woreʔ *make designs or decorations* → worekaŋ *be designed*.

There are a number of words in which the original consonant is retained before a fossilised suffix. It might appear at first that in these words the original consonant is recovered preceding suffixation. However, Maryott (personal communication) explains these forms:

There is a small class of words that may at first appear to have realisations other than the expected k or t. A more careful investigation reveals these forms to be artifacts of an obsolete, non-productive system explainable on historical rather than descriptive grounds. Some of these forms have counterparts, often with a shift in meaning, in the productive system.

His examples include San, Snl katatikilaŋ *sleeping loft, bed* (a form reflecting earlier *tikil *sleep* with fossilised affixation) besides katatikitaŋ *any place on which one sleeps* (a derivative of modern tikiʔ) and San, Snl sədapeŋ *west, place where sun sets* (reflecting earlier *sədap *to set (of sun)*, *go down* with fossilised suffix) besides sədakeŋ *be put down into something* (modern sədaʔ + -eŋ). Other examples are San gəliraŋ *be presented* (reflecting earlier *gəlid *give* + *-an) as well as gəlikaŋ *be presented* (a derivative of modern gəliʔ) and kakinaseŋ *small plate from which fish or meat is eaten* (reflecting earlier *kinaš *fish* + *-en), cf. modern kinaʔ.

4.2 Paragoge addition

The paragoge occurs in all languages except Rth, although it differs somewhat from language to language.

In Tal the paragogic vowel is a, occurring on all words which previously ended in a consonant other than *ʔ. (The original nature of the paragoge is discussed in the appendix.) The consonant preceding the paragoge is usually doubled, e.g. atuppa *roof* (PSan *atup), inassa *fish* (PSan *kinas), laŋitta *sky* (PSan *laŋit). It is not geminate if the preceding syllable contains a geminate consonant, e.g. annuma *six* (PSan *ənum), a nasal-stop cluster, e.g. sandaka *lean* (PSan *sander) or z, e.g. uzasa *wash* (PSan *uRas).

The paragoge does not occur on words which earlier had final *ʔ, either because *ʔ resisted the addition of the paragoge, as it did in Pre-San and Ban, or because PA developed subsequent to the historical loss of *ʔ. Where the paragoge now occurs after ʔ in Tal, ʔ reflects an earlier *k, e.g. utaʔa *hair* < Pre-Tal *utakka (PSan *utak), zusuʔa *rib* < Pre-Tal *zusukka (PSan *Rusuk).¹¹

In San and Snl the paragoge is əʔ: San uhasəʔ, Snl urasəʔ *wash* (PSan *uRas); San, Snl likudəʔ *back* (PSan *likud); San linuhəʔ, Snl rinurəʔ *earthquake* (PSan *linuR); San əhabəʔ, Snl rabəʔ *sharpen* (PSan *əRab); San beŋeləʔ, Snl weŋeləʔ *deaf* (PSan *beŋel).

In the Thl dialect of San the paragoge is iʔ, e.g. likuriʔ *back* (cf. Mang likudəʔ), Thl wiwihiʔ *lip* (cf. Mang biwiheʔ).

In Ban the paragoge is Vʔ, where V assimilates to the preceding vowel: uhasəʔ *wash* (PSan *uRas), dakeleʔ *many* (PSan *dakel), taŋisiʔ *cry* (PSan *taŋis), sokoloʔ *cough* (PSan *səkol), likuduʔ *back* (PSan *likud).

In present-day San, Snl and Ban the paragoge occurs on all words which would otherwise end in a consonant other than ʔ (whether original or the result of FCR) or a nasal.

In San, Snl, Ban and Tal the paragoge is lost if there is a suffix beginning with a vowel. Thus, San uhasəʔ, Snl urasəʔ, Ban uhasaʔ, Tal uzasa *wash* with the passive suffix become respectively uhasen, urasen, uhasen, uzasanna *be washed*.

5. AREAL DIFFUSION

Neither FCR nor PA occurred in a language ancestral to all the Sangiric languages; Tal does not reflect FCR and Rth does not reflect PA.

It is important to establish that neither FCR nor PA occurred in a language ancestral to San/Snl and Ban. Evidence is here given that Ban had split from the North Sangiric languages before either change appeared and consequently their occurrence in both Ban and San/Snl must be the result of areal spread.

The North Sangiric languages reflect metathesis of word-final *s and a preceding *t. In Pre-San the resulting final *t later reduced to ʔ. Metathesis was regular where *t was the consonant in the syllable immediately preceding final *s, as in the first two examples below. The change sometimes also occurred when *t was separated from final *s by two syllables, most examples recorded being of two-syllable words, with metathesis of initial *t and final *s, as in the second two examples below. This metathesis (MET) did not occur in Ban and Rth:

PSan	San	Snl	Tal	Ban	Rth	
*bitis	bisi?	wisi?	bisitta	bitisi?	witis	<i>calf</i>
*Ratus	hasu?	rasu?	zasutta	hatusu?	-	<i>hundred</i>
*taŋis	saŋi?	saŋi?	saŋitta	taŋisi?	taŋis	<i>cry</i>
*tages	sage?	sahe?	sahatta	tagese?	tahis	<i>reef</i>

MET is one of the strongest pieces of phonological evidence available for subgrouping the Sangiric languages, showing that San, Snl and Tal share a parent language not ancestral to Ban and Rth.

As Tal does not reflect FCR it must have split from San/Snl before FCR occurred. Since s-t metathesis is reflected in all three North Sangiric languages it follows that the occurrence of MET predated FCR (in fact, MET could not have occurred if final *t had already reduced to ?).

Since Ban does not reflect MET it, and Rth, had split from the North Sangiric languages before MET, and consequently before FCR, had occurred. Thus MET provides the important information that FCR did not occur in a language ancestral to Pre-San and Ban but that the innovation spread geographically subsequent to their period of common development.

The occurrence of PA in Ban and San/Snl must also be the result of areal spread. First, Ban and Rth form a subgroup but PA does not occur in Rth. Therefore its occurrence in both Ban and San/Snl cannot be the result of direct inheritance from a period of common development. Secondly, as is shown below, PA must have appeared in Ban and Pre-San chronologically later than FCR, which was itself the result of areal diffusion.

Also the occurrence of PA in San/Snl and Tal must be the result of areal spread rather than a shared inheritance from PNSan. As FCR occurred in Pre-San but not in Tal then it developed after the language split. Consequently, since PA appeared in Pre-San later than FCR, it also must have developed after the languages split.

Although the place of origin of the two innovations cannot be established some assumptions can be made. Barrack (1978:5) writes: "If we find evidence of lexical diffusion" [where an innovation is spreading through the lexicons of a number of contiguous dialects] "then we should expect the innovation to appear with greatest consistency in that dialect closest to the point of origin." He points out that there is no necessary reason why this should be so but that "the expectation of a higher degree of diffusion in dialects closest to the point of origin is corroborated by empirical investigation."

PA affected the whole of the Tal lexicon. It continued to operate in Ban and San/Snl after FCR and displaced it as the method for removing consonants from word-final position. It is unlikely that either of these languages would have displaced one successful method for dealing with final consonants by another innovation unless under external influence. Thus it is most probable PA began in the Tal-speaking area.

FCR clearly did not first appear in the Tal-speaking area as it did not operate there. FCR spread through more of the lexicon in San/Snl than in Ban, operating on consonants other than voiceless stops, and thus probably first began in the San/Snl speech area.

6. LEXICAL DIFFUSION

Chen (1972:468-469) writes:

Sound change does not operate on the lexicon en bloc and instantaneously or according to a uniform schedule; rather, it spreads itself gradually across the lexicon, and operates on words or groups thereof one after another ... this gradual spread of phonological change from morpheme to morpheme has become known under the name of 'lexical diffusion'.

He also states (p.468):

A phonological change can gradually extend its domain by extending its phonological environment. Thus a narrowly conditioned sound change with the initial shape as (1) may extend its phonological condition successively as (2,3...) and finally become an unconditioned change (n):

- (1) $X \rightarrow Y/C_1$
- (2) $X \rightarrow Y/C_{1,2}$
- (3) $X \rightarrow Y/C_{1,2,3}$
- ⋮
- (n) $X \rightarrow Y$

It sometimes happens that as a phonological rule diffuses through the lexicon, another rule appears which begins to operate on the same items, blocking the spread of the first rule. Chen and Wang (1975:256) write:

A phonological innovation may turn out to be ultimately regular, i.e. to affect all relevant lexical items, given the time to complete its course. But more often than linguists have thought, a phonological rule peters out towards the end of its life span, or is thwarted by another rule competing for the same lexemes.

The hypothesis put forward here is that FCR and PA were competing methods by which final consonants were dealt with in the Sangiric languages. According to this hypothesis the two processes began in different localities and spread through the contiguous speech communities (areal diffusion) and gradually through the lexicon in various localities (lexical diffusion).

In Pre-San and Ban FCR was the first rule to affect final consonants, operating on voiceless stops. All final voiceless stops underwent FCR in Ban and possibly also in Pre-San, although a few may not have, as is further discussed below.

It is possible that FCR originally operated only on *t. If t-reduction in Rth was not a separate development then this was certainly the case:

*t > ? / __#

However, this cannot be established and the rule in any case extended to the other voiceless stops in Ban and Pre-San. This is referred to as phase 1 in the operation of FCR:

C > ? / __#

FCR Phase 1: [+stop]
 [-voice]

When the process of PA reached Ban final voiceless stops, but no other sounds, had undergone FCR. PA then operated on all remaining final oral consonants (examples given below, this paragraph). But in Pre-San almost every other final oral consonant was affected by FCR in some words. The only consonant for which no examples have been recorded for the Pre-San period is b. Although b-reduction did occur in some words there are dialect differences and changes affecting this sound may have been late (see Section 7). Examples of the reduction of other consonants are: San, Snl gəli? *give*, cf. Ban gilidi?, Rth hilir (PSan *gəlid); San həbo? *slip into hole or mud*, cf. Rth uwoh (PSan *Rəbog); San rene? *rotten*, cf. Tal zenessa, Rth enes (PSan *Renes); San, Snl apu? *lime*, cf. Tal apukka, Ban apuhu? (PSan *apuR); San kuļu? *breadfruit*, cf. Tal uļukka, Ban kuhuhu? (PSan *kuluR); San beŋko? *bent*, cf. Tal beŋkola, Ban beŋkolo?, Rth weŋkol (PSan *beŋkol); San, Snl tiki? *sleep*, cf. Tal ti?illa, Rth tikil (PSan *tikil).

It is possible that FCR extended its domain to some classes of sounds before others; for instance, *b may have been affected later than some other sounds. But this cannot be clearly determined and since FCR eventually affected all oral consonants it is suitable here to recognise a second, general phase in its operation:

FCR Phase 2: C > ? / ___#
[-nasal]

Thus FCR began to affect classes of consonants other than voiceless stops in Pre-San, but it was stopped before it could spread to more than a limited number of lexical items ending in such consonants.¹²

There is evidence that FCR and PA overlapped in time of operation in Pre-San, as discussed below. Chen (1972:478-479) talks of sound changes in overlapping time relation with each other. He illustrates this with two rules, R7 and R8:

R7 a → b / ___c

R8 c → d

and writes:

If ... some ac sequences emerged as bd and some as ad then the divergent developments of ac forms would reflect a stage in historical change where two rules (R7, R8) overlapped in time and were simultaneously applicable to the same set of lexical items, with the result that whereas some lexical items underwent the changes in the order of R7-8, some other lexical items followed the reverse order of events. In this latter case R8 does not destroy the phonological environment of the input to R7 in all instances, but in some only. This is so because neither the diachronic ordering of R7-8 nor its reverse, R8-7, can account for the coexistence of both ad and bd forms.

It is proposed here that such overlapping of FCR and PA is partly responsible for the present-day variation in San and Snl. Chen's R7 and R8 can be equated with FCR and PA respectively for San/Snl:

FCR: C > ? / ___ Ø #

PA: Ø > ə? / C ___#

where C is any consonant except a nasal or ?.¹³

The PA rule as expressed above is conditioned where Chen's R8 is not. The restriction on C is necessary since the PA rule could not operate on the output of the FCR rule, the sequence **ʔəʔ not being possible. Where PA applied first FCR could not operate on its output since the condition for the change to C, that it be in word-final position, no longer obtained.

This mutual interference of the rules applied of course only during the period of their overlapping. FCR was the first rule to operate in Pre-San, completing phase 1 (this is assumed for the moment). Soon after phase 2 began PA also began to operate. Thus the changes were in competition for the remaining eligible lexicon until FCR petered out, leaving PA as the only rule still operating on final consonants. This can be depicted as in Diagram 1:

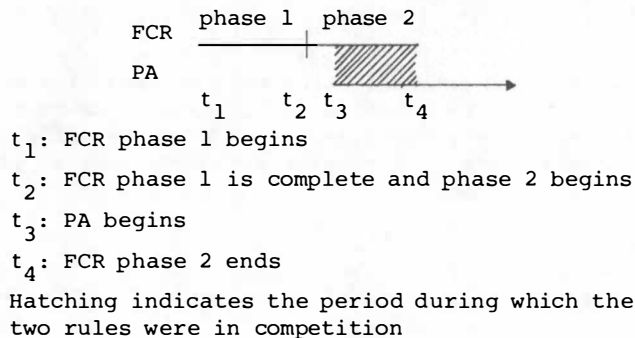


Diagram 1

This implies that PA became the 'favoured' method by which final consonants were dealt with in San/Snl, 'winning out' over FCR. Evidence for this comes from changes to borrowings in San and Snl. Almost all borrowed words ending in a consonant underwent PA; very few in which the final consonant underwent FCR have been identified. San, Snl ubaʔ *monkey* is apparently from a southern Mindanao language, where the word is widely distributed as uba|. Occurrence of b, for regular w, and the absence of cognates in the other Sangiric languages point to borrowing. It is possible that this was a very early borrowing, predating FCR and subsequently undergoing FCR when phase 2 of its operation began, as did a number of inherited words ending in *|.

There are only a few identified borrowings in which a voiceless stop reduced to ʔ in San: aŋkaʔ *strike up a song* from Malay (Mal) aŋkat, bebeʔ *duck* from Mal bebek, bianguʔ *beard*, with unexplained initial b, but Siau dialect dianguʔ from Mal jaŋgut.¹⁴ Northern San (Taruna, Tabukang) and Snl ətəʔ *winnow* is a borrowing (cf. Cotabato Manobo ətəp, Sarangani Manobo ʌtəp) beside Mang təəʔ and forms in the other Sangiric languages reflecting PSan *təəp.

Final consonants, both voiceless stops and other classes, reduced to ʔ in a few other borrowed words but all these words have doublets which retained the final consonant by means of PA (see Section 7).

With the exception of the few recorded cases referred to above, known borrowed words ending in a consonant took the paragoge in San.¹⁵ Most significantly this included voiceless stops. Borrowings which took the paragoge after a voiceless stop include: San harapəʔ, Snl ha|apəʔ *hope* from Mal harap; San sədapəʔ *delicious* from Mal sədap (cf. San sədaʔ *set (of sun)* < PSan *sədap); San kuatəʔ *strong* from Mal kuat; San umatəʔ *human* from Arabic via Mal umat; San porokəʔ *fork* from Dutch vork; San balakəʔ *beam* from Dutch balk; Snl utukəʔ *brain* from a Mindanao language (cf. Samal, Mansaka, Tagbanwa qutuk).

A number of words in San and Snl which have retained final p, t and k, taking PA instead of FCR, cannot be identified as borrowings. But the probability is that all such items were borrowed, entering the language after PA had replaced FCR as the method for dealing with final consonants. Of 112 items in Steller and Aebersold's dictionary ending in voiceless stop + əʔ (where groups such as kənditəʔ, kinditəʔ, kunditəʔ *thrifty* are counted only once) 31 can be identified as borrowings from Mal (often from Arabic or another language via Mal) and 35 from Dutch. Steller and Aebersold (1959) identify one as a borrowing from Chinese and one from Tidore. Of the remaining 44 items eight have vowel a in the environment əC__təʔ#, e.g. bakatəʔ *darkness*, kalatəʔ *curse*. It was a regular rule in PSan that PAN *a was replaced by *e before a final alveolar or dental consonant if the preceding vowel was also *a, separated from it by one consonant, e.g. PAN *baRat > PSan *baRet *wind*, PAN *Zalan > PSan *dalen *road*, PAN *palaj > PSan *paled *palm*. Thus these items are identified as borrowings by having a instead of e in the final syllable. Eight items have intervocalic voiced stops instead of corresponding continuants, b instead of w, d instead of r (l in Snl), g instead of ɣ (h in Snl), e.g. obotəʔ *pride*, udupəʔ *conscientious*, pagutəʔ *fastidious*. Voiced stops became continuants intervocalically in PNSan except after *ə and these eight items can be regarded as borrowings.¹⁶ One of them, uagatəʔ *kind of spirit; on quiet nights its loud steps can be heard outside*, contains both a instead of e and intervocalic g, both features identifying it as a borrowing (cf. the inherited form in San baheʔ *wind* < PAN *baRat).

Thus of the 112 items in Steller and Aebersold ending in voiceless stop + əʔ only 29 cannot be identified as borrowings on the available evidence. However, the fact that none of these has known cognates in the other Sangiric languages, apart from Snl, lends strong weight to the likelihood that they are also borrowings.

Nevertheless, the possibility must be left open that some of them are inherited from PSan. If this is so then the only explanation is that PA reached the Pre-San speech community before FCR had spread to the entire eligible lexicon in its first phase. In that case the period in which the two rules overlapped would have begun before phase 1 of FCR was complete so that there would have been a period in which the two rules were in competition for those lexical items still retaining final voiceless stops. If this were the case then Diagram 1 would have to be modified as in Diagram 2:

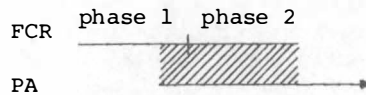


Diagram 2

Considering that PA became the preferred method for removing consonants from word-final position it seems unlikely phase 2 of FCR would even have begun if PA was already in operation. What seems more probable is that, if PA did overlap with phase 1 of FCR, the second, general phase of FCR began while phase 1 was still running its course (that is, consonants other than voiceless stops began to be replaced by glottal stop before this process had affected all voiceless stops). If so the overlap of the two rules can be represented as in Diagram 3:

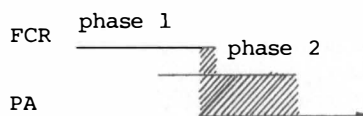


Diagram 3

Nevertheless, as mentioned above, the possibility is strong that all words ending in voiceless stop + paragoge were later borrowings and that phase 1 of FCR was completed before PA commenced in Pre-San, as represented in Diagram 1.

Diagram 1 shows PA continuing after FCR phase 2 had ceased to operate, evidence for this coming from borrowings, as mentioned above. There are no recorded cases of recent borrowings with final consonants undergoing FCR; all have undergone PA, which is thus still in operation. Examples of recent borrowings in San are: *pələtəʔ gramophone record* from Dutch (grammofoon-) *plaat*, *səpaltəʔ asphalt*, *tərakəʔ truck*, *listrikəʔ electric*, *konsoləʔ consul*, *motorəʔ motorboat* (information from K. Maryott, personal communication).

Although most, and probably all, words ending in a voiceless stop which took the paragoge in San were borrowed, many words ending in a consonant other than a voiceless stop which underwent FCR must have been inherited. It is important to establish this here because it could otherwise be argued that in San/Snl, as in Ban, all voiceless stops underwent FCR while all other consonants were retained through PA, the 'irregular' forms being treated as borrowings. Such an approach would have the advantage of imposing regularity of sound change on the language and the need to appeal to lexical diffusion would be obviated.

Charles (1974:463) refers to San *ləhaʔ pound rice again to get it white* as problematic because the expected form would be ***ləhasəʔ* (Proto-Philippine **DeRqas*). However, the evidence is convincing that some San words in which a previous final *s* was replaced by *ʔ* are directly inherited and need not be regarded as problematic. The list of PSan reconstructions in Sneddon 1984 contains 30 items with final **s* to which the paragoge was added in San but only three which underwent *s*-reduction and one case of doublets. The items undergoing FCR are San *heneʔ*, Snl *reneʔ rotten*, cf. Tal *zenessa*, Rth *enes* (PSan **Renes*); San, Snl *kinaʔ fish*, cf. Tal *inassa*, Ban *kinasaʔ*, Rth *kinas* (PSan **kinas*); San, Snl *nipiʔ thin*, cf. Ban *nipisiʔ*, Rth *nipis* (PSan **nipis*). San also has the pair *bəhiʔ*, *bəhisəʔ line, stripe*, cf. Ban *bihisiʔ* (PSan **bəRis*), which are treated in Section 7. (Other words which underwent *s*-reduction in San, such as *ləhaʔ* mentioned above, have no known cognates in South Sangiric languages and consequently have not been assigned PSan etyma.)

Although only a small percentage of inherited words with final **s* underwent FCR in San there are objections to any assumption that these were actually borrowings. First, with one known exception in San (*ibələʔ devil* from Mal (Arabic) *iblis*) and one in Snl (*kalətəʔ money* from Mal (Arabic) *kərtas*), all identified borrowings with final *s* underwent PA, such as San *maləsəʔ lazy* from Mal *malas*, San *galəsəʔ glass* from Dutch *glas*. Even the two known exceptions have doublets in which *s* was preserved by PA (see Section 7). Thus the absence of a paragoge itself cannot be used as evidence that a word was borrowed.

The difficulty with assuming borrowing can be seen if we look at particular cases of *s*-reduction, for instance the item San, Snl *kinaʔ fish*. Related forms occur in the other Sangiric languages but have not been recorded for other languages which are geographically close (although cognates occur in some Borneo languages - R. Blust, personal communication); therefore borrowing from an

external source would be difficult to maintain. If the word was borrowed from another Sangiric language then there are several problems: (a) Since Ban *kinasa*?, Tal *inassa*, Rth *kinas* attest to PSan **kinas*, this item must have been replaced in Pre-San by another word, only to be borrowed again later, a not impossible but nevertheless unlikely situation. (b) If it was borrowed before FCR and PA operated it would have been borrowed from Ban or Tal as **kinas*. But there is no reason why such a borrowed word would undergo FCR while inherited words did not. (c) If it was borrowed after the operation of PA it would have been borrowed as a form with the paragoge. Further, the word *kakinaseŋ* *small plate from which one eats fish or meat*, with fossilised affixation, attests to the earlier occurrence of **kinas* in Pre-San.

The same arguments apply to the other words which have cognates in other Sangiric languages ending in *s*: hene? *rotten* and nipi? *thin*. Thus the application of FCR, instead of PA, to such words cannot be accounted for in terms of borrowing. It must therefore be recognised that these were directly inherited words, reflecting PSan **kinas*, **Renes* and **nipis*. It follows that some words with final **s* in Pre-San did undergo FCR rather than the more common PA.

The case of *s*-reduction has been used to show that FCR affected some words ending in consonants other than voiceless stops before PA blocked it from spreading further through the lexicon, a situation illustrated in Diagram 1 above. The same arguments can be applied to words ending in other consonants as well.

7. DIALECT DIFFERENCES AND DOUBLETS

The theory of lexical diffusion accounts for dialect differences within San/Snl in the treatment of final consonants. Hsieh (1977:168) points out that a sound change "will continue to proceed at a different speed and influence the lexical items in a different order in one dialect than in another."

In its second phase FCR affected certain lexical items in some San/Snl dialects but not in others. Thus at a certain time a lexical item underwent FCR in dialect A but not in dialect B. PA was prevented from affecting the item in dialect A by the earlier application there of FCR but it operated on the item in dialect B thereby in that dialect blocking FCR.

Some examples of dialect differences in San are: Siau *inti*?, Mang *intilə*? *stretch*; Siau *biu*?, Mang *biu|ə*? *despise*; Siau *ka|umbia*?, Mang *ka|umbiagə*? *twisted*; southern dialects *lələ*?, Mang *lələbə*? *bald patch*; Mang *tuma*?, southern dialects *tumadə*? *crush*; Mang *əmbə*?, Siau *əmbəhə*? *bleat*; Taruna *lewo*?, Mang *lewohə*? *young coconut*; Tabukang *səhi*?, Mang *səhidə*? *chicken lice*; Taruna, Tabukang *ulu*?, Mang *uluğə*? *massage by rubbing*; dialect (unspecified in Steller and Aebersöld) *əti*?, Mang *ətibə*? *back pain*.

Differences between San (Manganitu dialect) and Snl (Sarangani dialect) include: San *bisu|ə*?, Snl *wisu*? *boil*; San *dompo|ə*?, Snl *dompo*? *fuel a fire*;¹⁷ San *apidə*?, Snl *api*? *immediately after*; San *maririhə*?, Snl *madidi*? *yellow*; San *aha*? *clouds on horizon outlining land below*, Snl *arabə*? *cloud cap on mountain*; San *hunu*? *small fire*, Snl *la-runusə*? *bonfire*; San *kuni*?, Snl *kunidə*? *turmeric*; San *sia*? *yell, scream*, Snl *siadə*? *cry (of infant)*.

Such dialect variation may be quite limited. Steller and Aebersöld list only 19 cases of doublets where the members are identified as belonging to different dialects of San. One reason for this could be that the modern dialects began to emerge from Pre-San only when the changes discussed here were well under

way.¹⁸ If this was so then most items would be similarly reflected in all dialects of San/Snl and this does appear to be the case. Assuming that situation, Diagram 1 can be modified as in Diagram 4 to indicate the period of the emergence of modern dialects:

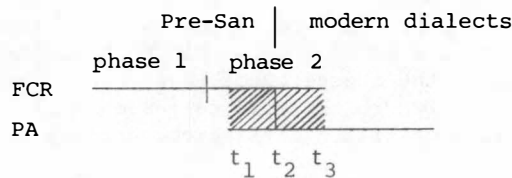


Diagram 4

Before t_1 only FCR was in operation; after overtaking all final voiceless stops it began to act on other classes of consonants. From t_1 until t_2 FCR and PA were in competition, each affecting some items ending in consonants; these items are reflected in the same way in all modern dialects of San/Snl. From t_2 to t_3 the competition continued but following different schedules in different dialects. It was during this period that the dialectally different forms emerged. At t_3 FCR ceased to operate; t_3 could of course have been a different real time in different dialects.

Steller and Aebersold list a number of doublets in San where one results from FCR and the other from PA and where the two forms are not identified as occurring in separate dialects. Examples are: *bəhi?*, *bəhiə?* *line, stripe*; *pongo?*, *pongo!ə?* *broken off*; *disi?*, *disihə?* *stand firm*; *kəmi?*, *kəmihə?* *silent*; *lau?*, *laugə?* *mix*; *həŋko?*, *həŋkodə?* *fear*; *gono?*, *gonobə?* *stalk*; *əlo toto?*, *əlo totobə?* *easily moved to tears*.¹⁹

The theory of lexical diffusion, which recognises the possibility of the overlapping in time of competing changes, can also offer an explanation for this variation. Hsieh (1977:163) points out that as a change spreads through the lexicon each item affected undergoes a period of synchronic variation between the old and new form.

At time t_1 the first item a_1 begins to change by acquiring an alternative form b_1 but at the same time maintaining the old form a_1 At time t_2 the old form a_1 in the synchronic variation is dropped and the new form b_1 is retained. At this time, another item a_2 acquires the synchronic variation $a_2 \wedge b_2$ "

We can imagine a situation where, for instance, **pongo!* *broken off* was affected by FCR. It underwent a period of synchronic variation **pongo!* \wedge **pongo?*. Then, before the next stage, loss of variant **pongo!* and retention of *pongo?* alone, could occur, PA operated on the variant **pongo!*, producing *pongo!ə?*.

Doublets occur among a few borrowed forms and they can likewise be explained, e.g. *pingu?*, *pingurə?* *thimble* from Dutch *vinger*; *sa!ado?*, *sa!odagə?* *slovenly* from Dutch *slordig*. In the case of *ibəla?*, *hibəlusə?* *devil* from Arabic via Mal *iblis*, phonological differences between the two forms suggest they may have been separately borrowed rather than the result of the above-described processes operating on a single borrowing.

Sometimes differences in meaning also suggest separate borrowing of doublets, e.g. San, Snl *tampa*? *place* as well as Snl *tampatə*? *tomb, shrine (of deceased royalty or holy man)* from Mal *tempat* and San *karətəsə*?, Snl *ka|ətəsə*? *paper* as well as Snl *ka|ətə*? *money* from Arabic via Mal *kertas*. In these cases the form with consonant reduction is likely to be the older borrowing.

Steller and Aebersold note four sets of doublets in San where a previous *t has undergone FCR in one member and PA in the other, two sets being borrowings from identifiable sources, *iŋa*? and *iŋatə*? *remember* from Mal *iŋat*, *pupu* (with irregular t-loss rather than reduction) and *puputə*? *crowbar* from Dutch *koevoet*, one set, *sahagi*? and *sahagitə*? *charm for accomplishing a job quickly*, almost certainly a borrowing because of irregular intervocalic g, though from an unknown source, and one set, *amba*? and *ambatə*? *adorn*, not identifiable as a borrowing.

Three items with earlier final voiceless stops are known where San and Snl reflect different processes (although semantic differences suggest two of these might not be directly related): San *kuku*?, San *sə-kukutə*? *kind of sea fish*; San *suri*? *recall dimly*, Snl *su|itə*? *learn*; San *bilu*? *oblique, slanting*, Snl *bilukə*? *to tack (of sailboat)*.

Although such variation involving final voiceless stops occurs in very few items it must be accounted for. The only explanations that can be offered at present are that competition between the two processes did operate marginally during the first phase of FCR, as in Diagram 3 (although if so far more evidence than this would be expected) or that all the items are borrowings, as seems likely, and that they were borrowed during the second phase of FCR, that process competing with PA for these forms as it did for forms ending in other consonants.

8. CONCLUSION

The apparent irregularity in the way consonants were removed from word-final position in San presents difficulties for any explanation based on the traditional view that phonological change is instantaneous and knows no exception.

On the other hand, the theory of lexical diffusion can account for the development of two different changes, FCR and PA, in the same phonological environments.

The occurrence of doublets in San, where one is the result of FCR and the other of PA operating on a common parent form, cannot be accounted for by traditional explanations of diachronic sound change. Borrowing, so frequently appealed to in the case of difficult phonological problems, offers no explanation since both clearly inherited forms with cognates in the other Sangiric languages and forms which are positively identified as borrowings have resulted in such doublets.

Nor can dialect mixture be seriously considered. Although Steller and Aebersold do not always identify dialect differences, sufficient information is available to show, as illustrated in Section 7, that dialect variation is random. The phenomenon of doublets which occurs in Mang is almost certainly a feature of the other dialects as well.

The theory of lexical diffusion offers an explanation for the origin of such pairs by recognising the possibility of one rule affecting a word while it was undergoing a period of synchronic variation between the original form and the form produced by the other rule.

The evidence presented above supports the hypothesis that FCR first operated in Pre-San but before it could spread to the entire eligible lexicon it was blocked by PA. FCR apparently did not come to a sudden stop with the arrival of PA to the San/Snl speech area but rather there was a period of competitive overlap, as represented in Diagram 4, the occurrence of doublets and dialect variation strongly supporting this.

Finally PA established itself as the preferred method for removing consonants from word-final position. PA is still in operation, with borrowings now undergoing PA rather than FCR to conform with the phonological pattern of the language.

The theory of lexical diffusion accounts for forms like San/Snl *nipi?* *thin*, for expected ***nipisə?* (PAN **nipsis*), and *apu?* *lime*, for expected ***apuhə?* (PAN **qapuR*). Some Austronesianists, using evidence from San in comparative studies, have regarded some forms as problematic or as borrowings because of such variation. This paper, while offering supporting evidence for the theory of lexical diffusion, shows that some apparently irregular forms in San are in fact directly inherited and can be employed with confidence in comparative studies.

APPENDIX

The paragoge is phonologically different in Ban, Tal and San/Snl. However, it is argued here that it was originally **ə* in all languages.

In Tal the paragoge could not originally have been **a*, as it is today. Since it is most unlikely that geminate consonants occurred word finally it can be assumed that PA chronologically preceded the development of doubled consonants. If the paragoge were originally **a* there would be no way to explain why doubling occurred to the consonant preceding the paragoge, e.g. *laŋitta sky* < PSan **laŋit*, but not before a where this reflects PSan **a*, e.g. *mata eye* < PSan **mata*. Consequently the paragoge must have been a vowel other than **a*. The same argument can be used against its having been **e*, **i*, **o* or **u*. However, the vowel of the paragoge could well have been **ə*.

Schwa occurred in PSan in all but final syllables. PSan **ə* was replaced by *a* in Tal, the following consonant having become geminate: PSan **əpat* > Pre-Tal **əppata* > Tal *appata four*, PSan **bəli* > Pre-Tal **bəlli* > Tal *balli buy*. It can be established that consonant gemination occurred prior to changes to **ə* for the same reason that gemination occurred prior to the paragoge becoming *a*, namely, that gemination does not occur following a where it reflects PSan **a*.

In Ban **ə* no longer occurs, having regularly assimilated to the following vowel: PSan **səlet* > Ban *sele?* *insert*, PSan **bəka* > Ban *baka split*, PSan **ləno* > Ban *lono smooth*. The shape of the paragoge in Ban can be accounted for if we assume the same process of vowel assimilation applied to it, i.e. if the paragoge was originally **ə?* then *ə* assimilated to the closest, in this case preceding, vowel: PSan **kinas* > Pre-Ban **kinasə?* > Ban *kinasə?* *fish*, PSan **apuR* > Pre-Ban **apuhə?* > Ban *apuhu?* *lime*, PSan **kəmis* > Pre-Ban **kəmisə?* > Ban *kimisi?* *squeeze*.

There is one San dialect, Tahulandang, in which the paragoge is not *ə?* but *i?*, e.g. *tiŋari?* *correct* (cf. Mang *təŋadə?*), *tuŋuhi?* *egg* (cf. Mang *təluhə?*). In Thl, as in Tal and Ban, previous **ə* has been replaced by other vowels in all

positions (as in the two examples above). Thus in Thl also the vowel of the paragoge can be accounted for in terms of replacement of earlier *ə.

In all languages except Tal the paragoge ends with *ʔ*. In Tal final **ʔ* was lost, e.g. PSan *kento^ʔ > Tal ento *lɪmp*, PSan *Ramu^ʔ > Tal zamu *red*. It can therefore be taken that the paragoge originally ended in glottal stop which was later lost by regular rule in Tal.

Thus there is good evidence that the paragoge was originally *ə^ʔ in all languages. It would be natural that when the rule of PA spread from Tal (see Section 5) it was adopted in the other languages in the same phonetic shape, later variation occurring with the replacement of *ə in all positions in Tal, Ban and Thl.

NOTES

1. I wish to thank Robert Blust and David Zorc for their helpful comments on an earlier draft. I am also very grateful to Kenneth Maryott who patiently answered questions on Sangil, providing much-needed information on that language. I also express my thanks to Professor Peter Worsley, Department of Indonesian and Malayan Studies, The University of Sydney, who provided the facilities which allowed me to commence work on this paper during a brief visiting fellowship in his department.
2. Charles (1974) believes, on lexical evidence, that San lies outside the Philippine group. On the other hand, Walton (1979), also using lexical evidence, finds San to be a first-order branch of Southern Philippine, one of the two first-order branches of the Philippine group. In a painstaking comparative study Zorc (1986) presents strong lexical evidence for recognising the Sangiric languages as a subgroup of the Philippine or Eastern Hesperonesian languages.
3. Maryott has recently begun to refer to Sangir as Sangihé, representing [səŋihə^ʔ], the indigenous name in Manganitu and Tabukang dialects, and to Sangil as Sangiré, representing the indigenous name [səŋirə^ʔ]. Not only could this be confusing to linguists but the name [səŋirə^ʔ] is also used by some Sangirese to designate their own language, i.e. in dialects such as Taruna where r corresponds to Manganitu and Tabukang h. In this work the better known, and less confusing, names Sangir and Sangil are used.
4. A lexicostatistical comparison gives San (Manganitu) and Snl (Sarangani) a cognate percentage of 82. Walton (1979) finds them to share 90% of their basic vocabulary, basing his study on material in Reid 1971 for which the San list was drawn from the Tabukang dialect as spoken by immigrants in Mindanao.
5. Two instances, apart from those considered in this paper, are the replacement of all final nasals by ŋ in San, Snl and Ban, e.g. San ənuŋ, Snl, Ban nuŋ *sɪx* < PSan *ənum, and replacement of medial y by l in San and Tal, e.g. San kalu, Tal alu *wood* < PSan *kayu.
6. Danie (1981) contains maps showing the areas of San settlement in northern Minahasa.

7. Adriani (1911:4,5) notes that the Tal dialects on the remote northern islands of Nanusa and Miangas exhibit FCR rather than PA. He provides only a half dozen examples and he and Steller (1913:4) provide a few examples from the far north Essang dialect which suggest it too may have undergone FCR. But since all the Essang examples and most of those for the other two dialects involve reflexes of *R, which has diachronically undergone a number of unique changes in Tal, the position is far from clear. Until more information is available these dialects cannot be further considered here.
8. Rth has ala? *fetch* for expected **alap (PSan *alap) but no assumptions can be made on the basis of one known occurrence of p-reduction, which must be left 'unexplained'.
9. Five Ban words have been recorded which end in voiceless stop + paragoge. These are all borrowings: pehete? *bat* (borrowed from an adjacent Minahasan language, without cognates in the other Sangiric languages), kulata? *fungus* (borrowed from a Minahasan language or Malay, without Sangiric cognates except in Rth which also borrowed the word), bebeke? *duck* (borrowed from Malay), punggutu? *stunted* (with irregular g following ŋ, instead of regular k - see Sneddon 1984:47), uagata? *strong wind* (with irregular g instead of h and occurrence of a before t - see p.61 of text for a discussion of this change in the Sangiric languages - cf. PSan *baRet *west wind*). These forms undoubtedly entered the language after FCR had run its course and when PA was operating on all remaining final oral consonants.
10. Ban tiki *sleep* (PSan *tikil) may be an exception. But although San, Snl tiki? results from l-reduction the absence of glottal stop in Ban suggests independent loss of *l, not associated with FCR.
11. The change *k, *kk > ? occurred subsequent to PA, i.e. the paragoge was added to *k, not to ?. This is established by items such as ba?isa *tie in a bundle* < PSan *bəkis. The consonant before the paragoge was not doubled where the preceding consonant was doubled. As s in ba?isa is single the preceding consonant was double at the time of paragoge addition. Therefore the word was earlier *bakkisa. The geminate *kk blocked doubling of s but later became ?. Therefore the rule *k(k) > ? occurred later than paragoge addition. The replacement of *k by ? in Tal occurred in all positions in the word (since glottal stop is interpreted as non-phonemic in initial position the change there is regarded as *k > Ø) and is therefore not a case of final consonant reduction.
12. For the majority of words ending in ? in San and Snl it is not possible to determine the original final consonant. This can only be done in cases where the consonant is preserved before a fossilised suffix (see Section 4.1) or where external cognates have been discovered. Most words ending in ? whose original final consonants have been identified are fairly common words with known cognates and it is certain there are others, as yet unidentified, which earlier ended in a consonant other than a voiceless stop. Since only a small number of basic vocabulary items previously ending with oral consonants other than voiceless stops underwent FCR it is very likely that only a very limited percentage of the entire eligible lexicon was so affected. In the list of PSan reconstructions in Sneddon 1984 (which includes only items with known reflexes in Rth and/or Ban as well as in North Sangiric languages) there are 128 items ending in oral consonants other than voiceless stops which have San reflexes which took the paragoge and only 16 with San reflexes undergoing FCR, as well as three doublets (which are discussed in Section 7).

13. The FCR rule would not normally be written with Ø before #. This is done here so that the symbol Ø can be used in both rules to highlight the similarity of the processes involved here to those described by Chen.
14. It is possible that borrowings from Mal with ʔ for expected t in fact come from a Mal dialect where final t had been replaced by ʔ or where, as in Manado Malay, final *t was lost, cf. Manado Mal *aŋka lift* (Mal *aŋkat*), *tampa place* (Mal *təmpat*), *iŋa remember* (Mal *iŋat*). San often added final ʔ to borrowed words ending in a vowel (see Sneddon 1984:52) and San words such as *aŋkaʔ*, *tampaʔ* (see text p.65) and *iŋaʔ* (see text p.65) could thus be borrowings from Manado Mal, as are *neneʔ grandmother* (Manado Mal *nene*) and *teteʔ grandfather* (Manado Mal *tete*).
15. Information on Snl is far from complete and no suggestion is made that other forms do not occur in that language.
16. San and Snl sometimes retained voiced stops morpheme initially after a fossilised prefix, e.g. San *kadadəməhəʔ*, Snl *kadadəməməʔ evening star* (from a root *dəməR), though not regularly. In Siau dialect (for which available information is insufficient to allow a phonological study) voiced stops sometimes occur intervocalically in inherited words, e.g. *laboʔ big*, cf. Mang *lawoʔ many*.
17. In Snl *l was lost preceding the paragoge and following a back or low vowel. The vowel preceding *l was also lost if unstressed, e.g. Snl *pundəʔ paddle* < PSan *pundəl, Snl *kapəʔ ship* < earlier *kapələʔ, cf. San *kapələʔ* (from Mal *kapal*). Such post-PA l-loss would have resulted in Snl **wīsəʔ, **dompəʔ. The forms *wisuʔ* and *dompoʔ* thus result from the final *l having undergone FCR.
18. Another reason for apparently limited dialect variation is incompleteness of information in Steller and Aebersold. This is essentially a dictionary of the Manganitu dialect and information on other dialects is only irregularly given.
19. Steller and Aebersold give 68 doublets which have either identical meanings (48) or very similar meanings (20). Borrowed or probably borrowed forms are not counted here, nor are pairs whose meanings are not at least very similar. It cannot automatically be assumed that where there are such doublets the one with final ʔ represents reduction of the same final consonant as occurs in the other. There are numerous examples in San of doublets with different final consonants, e.g. *əŋgehəʔ*, *əŋgesəʔ noise of flowing water*; *latubəʔ*, *latugəʔ joke*; *lokabəʔ*, *lokasəʔ bract*. There could well have been doublets with different final consonants where one underwent PA and the other FCR. If we look at Malay we see pairs such as *lətup*, *lətus explode*. If this pair had occurred in Pre-San the likely reflexes would be **lətuʔ and **lətusəʔ, giving the appearance of two forms reflecting one etymon. Steller and Aebersold sometimes cross-reference forms which are not from the one etymon, e.g. *onaʔ fishscales* and *onəsəʔ cuttings, peelings*, where the former actually reflects PSan *onap. Nevertheless, considering the number of doublets the probability is that there are many which reflect a single etymon which underwent both FCR and PA according to the process described here; for instance *səluʔ river* and *səluhəʔ gutter, riverbed* both clearly reflect PAN *saluR *waters*.

REFERENCES

- ADRIANI, N.
 1893 *Sangireesche spraakkunst*. Leiden: Nederlandsch Bijbelgenootschap.
 1911 *Een en ander over het Talaoetsch*. The Hague: M. Nijhoff.
- BARRACK, Charles M.
 1978 The High German consonant shift: monogenetic or polygenetic?
Lingua 44:1-48.
- CHARLES, Mathew
 1974 Problems in the reconstruction of Proto-Philippine phonology and
 the subgrouping of the Philippine languages. *Oceanic Linguistics*
 13/1&2:457-509.
- CHEN, Matthew Y.
 1972 The time dimension: contribution toward a theory of sound change.
Foundations of Language 8:457-498.
- CHEN, Matthew Y. and William S.-Y. WANG
 1975 Sound change: actuation and implementation. *Language* 51/2:255-281.
- DANIE, J. Akun
 1981 Peta bahasa di Minahasa bahagian timur laut. Paper presented at
 the Third International Conference on Austronesian Linguistics, Bali.
- HSIEH, Hsin-I.
 1977 A new method of dialect subgrouping. In William S.-Y. Wang, ed.
The lexicon in phonological change. The Hague: Mouton.
- MARYOTT, Kenneth R.
 1977 The phonemes of Sarangani Sangiré. *Studies in Philippine Linguistics*
 1/2:264-279.
 1978 Pre-Sangir *l, *d, *r and associated phonemes. *Studies in Philippine*
Linguistics 2/1:115-138.
- MILLS, Roger F.
 1975 Proto South Sulawesi and Proto Austronesian phonology. Ph.D.
 dissertation, University of Michigan (2 vols).
 1981 Additional addenda. In Robert A. Blust, ed. *Historical linguistics*
in Indonesia, 59-82. NUSA 10, Jakarta.
- PEI, Mario
 1966 *Glossary of linguistic terminology*. New York: Anchor Books.
- REID, Lawrence A., ed.
 1971 *Philippine minor languages: word lists and phonologies*. *Oceanic*
Linguistics Special Publication 8. Honolulu: University of Hawaii
 Press.
- SNEDDON, J.N.
 1984 *Proto-Sangiric and the Sangiric languages*. PL, B-91.

STELLER, K.G.F.

1913 *Nadere bijdrage tot de kennis van het Talaoetsch.* The Hague: M. Nijhoff.

STELLER, K.G.F. and W.E. AEBERSOLD

1959 *Sangirees-Nederlands woordenboek met Nederlands-Sangirees register.* The Hague: M. Nijhoff.

TEAM FAKULTAS SASTRA UNSRAT

1976-77 *Struktur bahasa Talaud.* Manado: Departemen Pendidikan dan Kebudayaan.

WALTON, Charles

1979 A Philippine language tree. *Anthropological Linguistics* 21/2:70-98.

ZORC, R. David

1986 The genetic relationships of Philippine languages. In *FOCAL II*, 147-173.

THE SAMOSIR DIALECT OF TOBA-BATAK

J.P. Sarumpaet

1. INTRODUCTION¹

Samosir is an island on beautiful Lake Toba in North Sumatra. It is 520sq. km in area, covering almost half of the 1300sq.km of the lake, and is situated 2°35'N, 98°48'E. The length of the island is about 43km and the widest part is about 20km. The highest peak on the ridge that runs in the middle of the island rises to 1,630m, whereas the lake itself is 911m above sea level.

Administratively, Samosir covers five subregencies (kecamatan), four on the island and one on the mainland.² On the island we have the subregencies of Pangururan, Simanindo, Onanrunggu and Palipi, whereas on the mainland we have the subregency of Harian Boho (see Map). The population is 126,696, grouped in 21,396 households (Hutajulu 1982:30).

Just west of the principal town of Pangururan is Mount Pusukbuhit (2,005m), at the western foot of which are the villages of Sagala and Limbong. According to tradition, the secluded and secure valley between Limbong and Sagala was called Sianjur Sagala Limbong Mulana. It was here that the first Batak settlement was established in the distant past, and it was from here that people dispersed over the centuries to form the various Batak ethnic subgroups of Karo, Pakpak-Dairi, Simalungun, Toba and Angkola-Mandailing (Joustra 1910:25; Hoetagoeng 1926:22; Parlindungan 1964:19; Parkin 1978:10-11; Sinaga 1981:189-224).

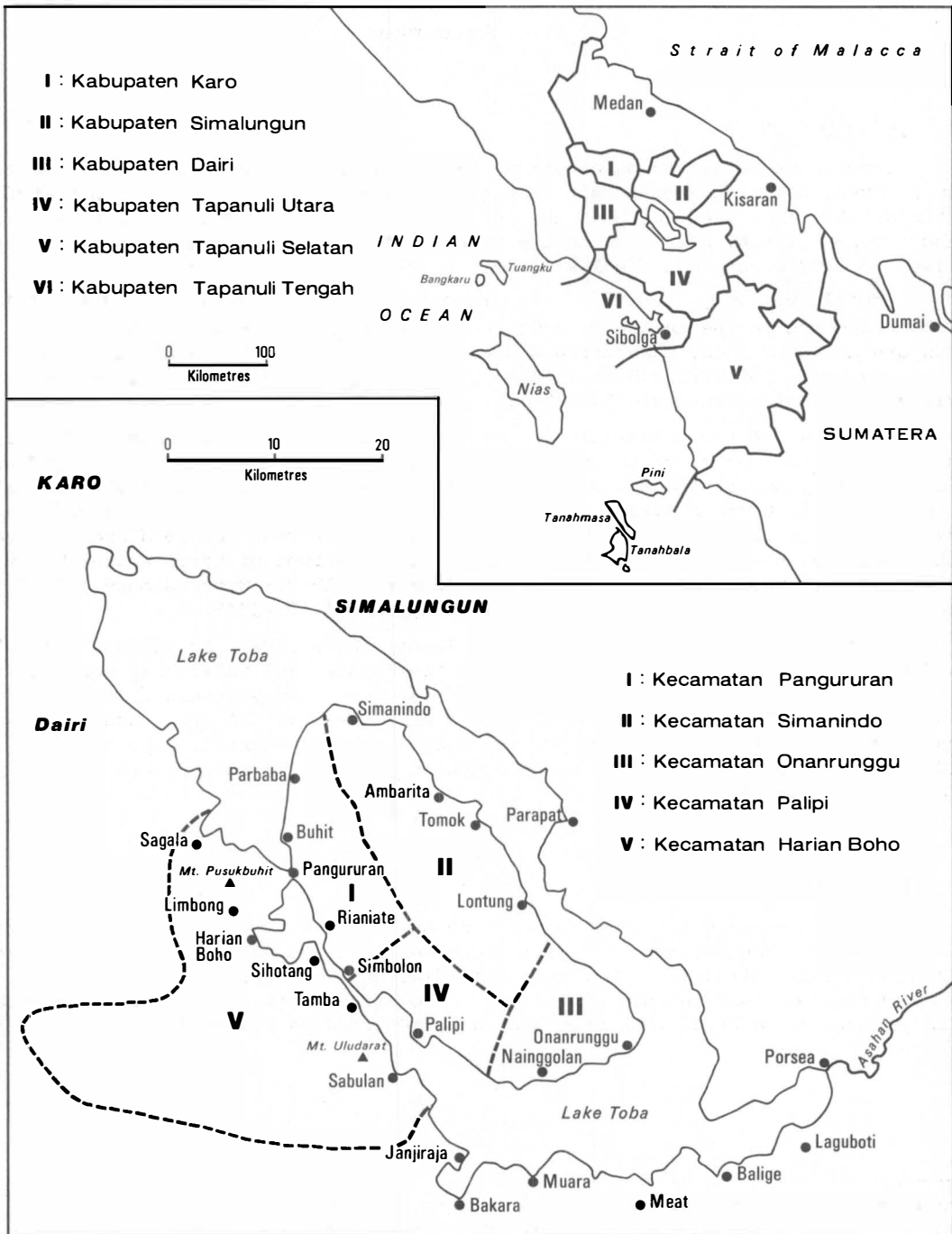
From time immemorial, Lake Toba and Samosir have always occupied a magical and revered place in the belief system of the Bataks, and until they were first sighted by Europeans in the 19th century, they were also a fascination for 'the white-eyed people' (si bontar mata) from across the seas. A great debate arose as to who the first European was to lay eyes on this mysterious lake (Pleyte 1895). Lake Toba was and to a great extent still is 'the sacred lake of the Bataks' (Braasem 1951).

2. THE BATAK LANGUAGES

It was customary to refer to the languages spoken by the Bataks of Sumatra as 'dialects' (Van der Tuuk 1864:iv; Schreiber 1874:8; Joustra 1910:10; Voorhoeve 1955:14; Nababan 1981:1). They were the 'dialects' of Karo, Pakpak-Dairi, Simalungun, Toba and Angkola-Mandailing. The differences in phonology and vocabulary, however, dictate that they should be regarded as individual languages.

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 73-86. *Pacific Linguistics*, C-94, 1986.

J.P. Sarumpaet



North Sumatra, with the Samosir dialect area

Hence the current custom - as is also followed by this writer - of naming them respectively Karo-Batak, Pakpak and Dairi-Batak, Simalungun-Batak, Toba-Batak, and Angkola and Mandailing-Batak.³

The linguistic divisions roughly coincide with the current administrative divisions of the area: Karo-Batak is primarily spoken in the Karo regency (kabupaten Karo), Pakpak and Dairi-Batak in the Dairi regency, Simalungun-Batak in the Simalungun regency, Toba-Batak in the North-Tapanuli (Tapanuli Utara) regency and Angkola and Mandailing-Batak in the South-Tapanuli (Tapanuli Selatan) regency (see Map).⁴

3. THE TOBA-BATAK LANGUAGE

The Toba-Batak language, then, is primarily spoken in the regency of North Tapanuli. North Tapanuli has as its capital Tarutung, which is situated about 48 kilometres from Balige. It is in this regency that one can hear Toba-Batak spoken in its less impure form.

Throughout Indonesia, wherever a Toba-Batak settles at the present time,⁵ he may still speak Toba-Batak at least to his wife, if she is also a Batak. Depending on circumstances, such as the number of Toba-Bataks living in his immediate environment, or the availability of services conducted in that language in his local church,⁶ he may also on occasion speak that language to his children. Over the years, however, as his children grow up in an Indonesian, rather than a Toba-Batak environment, he may have to be contented with the use of Indonesian (or in the case of a Jakarta resident, of Jakartese, and that one not too pure either) with his family.⁷ For obvious reasons, a Toba-Batak male married to a non-Toba-Batak female would normally speak Indonesian at home.

4. THE SAMOSIR DIALECT

Within Toba-Batak, we can distinguish one geographical dialect, namely the Samosir Dialect. There are other dialects distinguished by speakers of the language, but the differences are mainly slight differences in phonology.⁸

In this paper, I propose to discuss the characteristics of the Samosir Dialect (SAM). As the data show, the difference between it and Standard Toba-Batak (STB) is almost entirely lexical. There is hardly any syntactical difference between the two.

The material upon which this paper is based comprises recordings made in the field in late 1982, complemented by a few items which have appeared in written form in some dictionaries.⁹

The recordings consist of the speech of people who lived in out-of-the way villages, whether on the mainland side of the lake, or on Samosir Island itself. They are not elicited responses, but natural conversations in markets, jetties and eating places. The speakers were mostly women, between 25 and 50 years of age, who at most would have finished third grade in the local village primary school. Thus their speech was less adulterated by Standard Toba-Batak as spoken by those with more schooling, nor by Indonesian as spoken by those who had travelled farther away from Samosir.

It is interesting to note in the recordings how a conversation was conducted by three persons with different backgrounds. The first person lived in one of the villages in South Samosir and spoke the less adulterated Samosir Dialect. This means to say that she used most of the lexical items listed in the Appendix below. The second person had lived in the subregency capital of Pangururan all her life and had also finished Junior Secondary School. She spoke more or less standard Toba-Batak. The third person was a younger one who had been to Jakarta to visit relatives and who had stayed and worked in the provincial capital, Medan, for a year or two before returning to her native area. This person's speech was quite a mixture of Indonesian and Standard Toba-Batak; the Samosir Dialect was not represented at all in the mixture!

5. MARGINALITY AND ISOLATION

A look at the geographical position of Samosir shows that the language community we are concerned with here is situated right at the top margin of the Toba-Batak linguistic area: bounded to the north and north-east by the Simalungun language, to the north-west by the Karo language and to the west by the Pakpak-Dairi language.

Samosir was the last section of the Toba-Batak area to enter the mainstream of political and cultural life in Indonesia throughout its history. It was the last to be conquered by the Dutch colonial government, and it was also the area into which non-traditional religion, i.e. Christianity or Islam, came last. Even today, the Huria Kristen Batak Protestan, the largest of the Batak churches, still has one department in its central board specifically concerned with missionary work among those people in Samosir who still embrace the old traditional religion. The rest of North Tapanuli has not needed missionary work among followers of the old religion since the beginning of this century.¹⁰

6. NORTH AND SOUTH SAMOSIR

Samosir can be divided into North and South Samosir by drawing a straight line from Pangururan on the west coast to Lontung on the east coast as an imaginary boundary (see Map). Linguistically, this line is also significant. The data show that North Samosir is indeed different from South Samosir, with Pangururan as the largest town on the island occupying a rather marginal position. The villages in the subregency of Harian Boho have the linguistic characteristics of South Samosir.

Contrary to the predictions of some informants,¹¹ the data show that the Samosir Dialect is the same in both North and South Samosir, except for a few lexical items which are asterisked in the Appendix below. Indeed, fewer people in North Samosir speak the Samosir Dialect, and those who do, speak a more adulterated form of it. The exposed position of North Samosir to the outside world is certainly a determining factor here. Note, for example, the words *atas* 'above' and *kansang peanut*, which do not occur in our data from South Samosir. The first must have come from Malay into everyday usage in North Samosir, although the word was apparently widely used in Toba-Batak last century (Schreiber 1874:8). Today, it is more often noticed in Toba-Batak proper names, as in *Siatasbarita*, a mountain ridge near Tarutung, and in *Haunatas* (< *Hau na atas*), the name of a village east of Laguboti. *Kansang* would be very difficult to pronounce for inhabitants

of remote villages, as originally, Toba-Batak has no initial k and this phoneme is represented by ha in the Toba-Batak script. The word *kansang* is now common among urbanised young people when speaking Toba-Batak. *hu jai* and *huju ai* are borrowings from the neighbouring Simalungun language.

In South Samosir itself, it was noticeable that people from the mainland villages of Sihotang, Tamba and Sabulan spoke a purer form of the Samosir Dialect than those living across on the island. Isolation is certainly a determining factor here, as the villagers, in order to venture into the wider world of Sumatra and beyond, can only go through Pangururan, Balige or Porsea. Passage to the west, if it exists at all, is only in the form of narrow footpaths up the steep mountain range which has Mount Uludarat (2172m) as the highest peak.

Thus the Samosir Dialect is primarily spoken south of Pangururan (and to a certain extent around Pangururan itself): in Rianiate, Simbolon, Palipi, Nainggolan and Onanrunggu - to mention just the larger villages - on the south-western and south coast of the island, and also in the villages across from there on the mainland, namely the villages of Sagala, Limbong, Harian Boho, Sihotang, Tamba and Sabulan.¹²

It would be safe to speculate that since the only access to the mainland villages of Sihotang, Tamba and Sabulan is by crossing the lake, the Samosir Dialect will remain in use longer there.

7. OLD BATAK, THE SAMOSIR DIALECT (SAM) AND STANDARD TOBA-BATAK (STB)

If we accept the tradition that the Samosir area, especially the valley to the west of Mount Pusukbuhit, was the original settlement of the Batak, we can reasonably assume that SAM forms are older than STB forms. One might even speculate that the SAM vocabulary of today - except for the obvious loanwords from Malay or Indonesian such as *kaling* - contains remnants of Old Batak, the parent language of all the Batak languages of today.

The following characteristics can be seen in the Samosir Dialect when compared with Standard Toba-Batak. The standard orthography is used here. (See Appendix below for the complete data and for a guide to pronunciation.)

1. Zero Initial Consonant in SAM

The consonants missing in SAM are d, t, n and p.

The most frequent occurrence is the non-existence of the consonant d in SAM for both the preposition *di* *for; at; with* and the passive prefix *di-*: SAM *i imana* : STB *di ibana for him/her*; SAM *i au* : STB *di ahu with me; for me*; SAM *i dia* : STB *di dia where?*; SAM *i si* : STB *di si there*; SAM *i on* : STB *di son here*; SAM *i allang* : STB *diallang eaten*; SAM *i pingkir* : STB *dipingkir thought*; SAM *i lean* : STB *dilean given*; SAM *i tingkir*; STB *ditingkir peeped at; seen*.

We also have SAM *ambai* (now also a variant in STB): STB *tambai to increase*; SAM *asida* : STB *nasida they*; SAM *i allangna* : STB *ni allangna eaten by him*; SAM *inna* : STB *ninna he/she says*; SAM *unga* : STB *nunga*; SAM *intor* : STB *pintor immediately*.¹³

2. Zero Medial Consonant in SAM

The missing medial consonants are d, j and r.

SAM aong : STB adong *there is*; SAM jai : STB jadi *so; therefore*; SAM ni on : STB ndi on *this one*; SAM gijang : STB ginjang (pronounced gijjan) *above*; SAM nai : STB nari *still; more*.

3. SAM h : STB t

SAM hu : STB tu *to; towards*; SAM hu bagas : STB tu bagasan *towards the inside*; SAM hu on : STB tu son *this way*; SAM hu sai : STB tu sadu *that way*; SAM hu san : STB tu san *that way*; SAM hu toru : STB tu toru *downstairs; downward*; SAM hu gijang : STB tu ginjang *upwards*.

SAM huju must be paired with STB tuju *aiming at* as in SAM huju an : STB dompak an *in that direction*.

4. Zero Suffix in SAM

SAM bagas : STB bagasan *inside*.

5. SAM h : STB p¹³

SAM hainte : STB painte *wait!*

6. Short Form in SAM

SAM a : STB nunga *already*.

7. SAM k : STB t

SAM ake : STB ate *I say ...*; SAM eka : STB beta, which colloquially is often pronounced eta *let's go*.

8. SAM m : STB b

SAM imana : STB ibana *he; she*

9. SAM l : STB ng

SAM jagul : STB jagung *maize*

10. SAM i : STB e

SAM kaling : STB kaleng *tin; measure of about 16kg of rice*.

11. SAM r : STB d

SAM ndara : STB ndada *not*.

12. SAM e : STB d

SAM nei : STB ndi *this one here*.

13. SAM n : STB d

SAM nungkon ni i : STB dungkon ni i *apart from that; and then; by the way ...*

14. Final ng dropped in SAM

SAM onde : STB ondeng *the aforementioned*.

15. an infix in SAM

SAM hanami : STB hami *we(excl.)*.

16. Vowel Transposition in SAM

SAM haroa; horoa : STB huroha (colloquially pronounced huroa) *would it be?*;
 SAM de : STB do emphatic particle.

17. Consonant Transposition in SAM

SAM mana : STB nama *only*.

18. Contraction and Simple Forms in SAM

There is contraction in the SAM variants SAM huju ai > hu jai : STB tu san *that way*; i son > i on *here* : STB di son *here* and of SAM ni i > simply ni *of it*; *of that* and of SAM ni on > non. Simplification is evident in SAM lak : STB laos *just*; *simply* and SAM lok : STB loas *let it be*.

19. SAM t : STB r

SAM anta (pronounced atta): STB arta *property; wealth; possession*.

20. Longer form in SAM

SAM i sinon : STB di si *over there*; SAM musengani : STB muse *then; later on*.

8. FURTHER RESEARCH

The use of small regional languages is on the decline in Indonesia, mainly because it is too costly to mount television or even radio programmes using those languages. The school curriculum is also so crowded already, that the teaching of regional languages is almost impossible in many provinces. In the case of Toba-Batak, we can still expect it to continue to be used in church services - albeit in a diluted or contaminated form - at least in the Toba-Batak area.¹⁴

But what would be the chances for survival of a dialect like that of Samosir?

It is unfortunate that we have no data to show us the dialectal situation in Samosir in the past, to enable us to compare the data in this paper with them. Informants gave this writer anecdotal evidence that from the 1930s to the 1940s SAM was still very widely used.

It would be useful to conduct another investigation around 1992, to see to what extent the data in this paper have been lost to posterity.

NOTES

1. I am indebted to my principal informants, Mr D. Sarumpaet of Pematang Siantar, Mr P.A.S. Sibarani of Padang Sidempuan and Mr S. Sarumpaet of Jakarta for their invaluable help. They lived in various parts of Samosir before the war.
2. The island of Samosir was originally not an island, but a peninsula connected to the mainland by a low and swampy isthmus at the foot of Mount Pusukbuhit. Sailing vessels had to be dragged by men across the isthmus in order to

proceed north or south on the western strait. It became an island when a canal (henceforth called Tanoponggol) was constructed in 1906 on the isthmus to facilitate lake transport. This canal had almost silted up in the 1960s and the 1970s, and at the time this study was conducted, heavy earth-moving equipment was being used to widen and deepen it, as part of the provincial government's development project. The length of the canal is about one kilometre.

3. This system of referring to the Batak languages as languages rather than dialects is now the general practice in the academic departments of the tertiary institutions in North Sumatra, as indicated in the titles of the theses submitted to and passed by them.

However, if one defines dialects as languages which are mutually intelligible, one can certainly say that Toba-Batak and Angkola and Mandailing-Batak are actually two dialects of the same language.

4. The regency of Central Tapanuli (Tapanuli Tengah) is in a marginal position. Many Toba-Batak live there, as do many Angkola-Mandailing. There is also a Sibolga Dialect, which is a dialect of Malay, spoken there by the original inhabitants.

Some people make a clear distinction between Angkola and Mandailing, thus adding a sixth to the five ethnic subgroups. Linguistically at least, they can be regarded as one unit. See the title of Siregar 1977. The periodical *Bona Bulu*, published in Jakarta by an association of people from Angkola-Mandailing, now uses the term bahasa Tapanuli Selatan (the language of South Tapanuli).

5. Because Tapanuli is not very fertile, and its economy has always been rather backward, it has always been the ideal among Batak families to help their children to advance as far as possible in life through education. Anakkonhi do hamoraon di ahu *My children are my wealth* is a famous line from a very popular Batak song. This means that most Bataks have had to acquire education and make a living away from Tapanuli. Since the beginning of the century Bataks have migrated to Simalungun and East Sumatra (Cunningham 1958 passim). There was only one Junior Secondary School in the whole of Tapanuli, and only one Senior Secondary School in the whole of Sumatra before 1943. The latter was a five-year *Hogere Burger School* for Europeans, which accepted only a few Indonesians. Hence the great financial sacrifice on the part of parents to send their children to Java before the war.
6. In the early 1960s, the small Toba-Batak Christian community in Salatiga in Central Java decided that services be conducted alternately in Toba-Batak and Indonesian. The majority of the members were staff and students of Satya Wacana Christian University in that town, which included several from Nias! It is interesting (or sad ?) to note that by the late 1970s services were already completely Indonesian.
7. This inevitable code-mixing has also crept into the formal language used in sermons. In a forty-minute sermon in a church in Medan, thus not far from the centre of Toba-Batak culture, this writer noted the use of some 18 Indonesian words which would have been substituted with Toba-Batak terms by another preacher who was more careful or whose linguistic ability was greater. As it turned out, the preacher had his postgraduate training in the United States, and had just returned from a five-year stay in Europe, where he worked for one of the international church agencies.

8. Nababan (1981:1-2) mentioned the dialects - subdialects to him - of Toba Holbung, Humbang, Silindung and Hurlang.
9. The Samosir Dialect was usually treated by lexicographers of Toba-Batak (Hariara n.d.; Warneck 1977) on an equal footing with the languages of Pakpak and Dairi, Karo, Simalungun and Angkola and Mandailing, as can be seen in the way the Samosir Dialect was entered in their dictionaries. The actual number of lexical items in SAM, however, was always low.
10. See Siahaan 1982 and Hutabarat 1982 on the mass baptism of 59 people in February and another 88 people in March 1982 in the hills of the subregency of Pangururan. In the 1980 census, however, no followers of the traditional religion were reported in the regency of North Tapanuli. There were 21,620 Muslims, 21,862 Roman Catholics, 563,969 other Christians (i.e. Protestants), 173 Hindus and 378 Buddhists (Biro Pusat Statistik, Kantor Statistik Propinsi Sumatera Utara 1981:13).
11. Some informants who lived in South Samosir went as far as to say that people in North Samosir did not speak the Samosir Dialect at all.
12. Whereas Janjiraja is rather marginal, Bakara, Muara and Meat are right outside the Samosir Dialect area. The last three are accessible overland from the south-west, although villagers prefer to use lake transport to go to the markets of Balige, Pangururan and Porsea. There is one word which is peculiar to Muara, namely *tian* : STB *sian from*. This word was already noted by Van der Tuuk, and was found on the royal seal of Sisingamangaraja XII.
13. Van der Tuuk (1864:11) noted that some people living near Lake Toba were unable to pronounce the plosive p. In most cases, however, people would not have had difficulty in pronouncing very common words such as *pudi back*; *rear*; *api fire* and *gotap broken*, and indeed the name of the sacred Mount Pusukbukhit, for which there are no specific forms in the Samosir Dialect. Compare also with Intimate Toba-Batak in Sarumpaet 1982:22-33.
14. See, however, note 6 above. A motion was put in the 1981 meeting of the General Synod of the Huria Kristen Batak Protestan that the hymnal of the church be translated into Indonesian as a matter of urgency, because more and more churches had felt the need to conduct services in the Indonesian language rather than in Toba-Batak (Huria Kristen Batak Protestan 1981:269).

The periodical *Immanuel*, the official organ of the Huria Kristen Batak Protestan, which has appeared uninterrupted - except during the Japanese occupation - for 94 years, now also contains more and more articles written in Indonesian.

On the other hand, this writer is pleased to note that the language used in the lyrics of popular Batak songs is on the whole pure Standard Toba Batak, although the spelling in the leaflets accompanying the cassettes may make one shudder. These songs are most popular - often sung by non-Bataks who obviously must have gone through a lengthy period of elocution training - throughout Indonesia, and new songs continue to be written. We may be witnessing here a lively and more lasting genre of a regional literature.

APPENDIX

Pronunciation:

Medial consonant clusters, and those across word-boundaries, are pronounced as indicated between square brackets: mb [bb] as in ambai; mp [pp] as in jumpa; ngk [kk] as in tungkan and ndang koro; ngn [ŋŋ] as in iallangna; ngp [kp] as in ndang poro; nd [dd] as in onde and han dia; nj [jj] as in tinjang and han jolo; np [pp] as in han pudu; ns [ss] as in kansang; nt [tt] as in anta.

The n is silent in ndara whereas the ŋ is doubled in unga.

* = Lexical items peculiar to North Samosir.

Samosir Dialect	Standard Toba-Batak	English
a	nunga	<i>already</i>
abe	eda	<i>(for a woman) brother's wife; husband's sister</i>
age	agia	<i>whatever</i>
ai*	an	<i>that</i>
ni ai*	ndi an	<i>that one</i>
tungkan ai*	tungkan an	<i>a little bit that way</i>
ake?	ate?; atehe?	<i>I say ...</i>
alea!	ale!	<i>I say ...</i>
ambai	tambai	<i>to increase</i>
anta	arta	<i>property; wealth</i>
antirha	gadong hau	<i>cassava</i>
aong	adong	<i>there is</i>
asida	nasida	<i>they; he/she (polite)</i>
atas*	ginjang	<i>above</i>
bagas	bagasan	<i>inside</i>
... bo!	... ba!	<i>I say ...</i>
boras	parbue	<i>fruit</i>
bue; mamue	anak; maranak	<i>to have young (of animals)</i>
de	do	<i>(emphatic particle)</i>
duatna; duatni	duansa	<i>both</i>
eka	beta	<i>let's go</i>
eta	beta	<i>let's go</i>
gijang	ginjang	<i>above</i>
hainte	painte	<i>wait!</i>
han	sian	<i>from</i>
han dia	sian dia	<i>from where</i>
han jolo	sian jolo	<i>from the front</i>
han pudu	sian pudu	<i>from behind</i>
hanami	hami	<i>we (excl.)</i>
hanima	hamuna	<i>you (pl.)</i>
harbue	dahanon	<i>husked rice</i>
haroa	huroha	<i>would it be</i>
horoa	huroha	<i>would it be</i>
hu	tu	<i>to (prep.)</i>
hu atas*	tu ginjang	<i>upwards</i>
hu bagas	tu bagasan	<i>towards the inside</i>
hu gijang	tu ginjang	<i>upwards</i>

Samosir Dialect	Standard Toba-Batak	English
hu jai*	tu san	<i>that way</i>
hu on	tu son	<i>this way</i>
hu sai*	tu san	<i>that way</i>
hu san	tu san	<i>that way</i>
hu toru	tu toru	<i>downstairs; downward</i>
huju	dompok	<i>in the direction of</i>
huju ai*	dompok an	<i>that way</i>
huju an	dompok an	<i>that way</i>
i?	e?; beha?; i do?	(confirmatory question tag)
i	di	<i>for</i>
i imana	di ibana	<i>for him/her</i>
i	di	<i>at; with</i>
i au	di ahu	<i>with me; for me</i>
i dia	di dia	<i>where?</i>
i on	di son	<i>here</i>
i si	di si	<i>there</i>
i sinon	di si	<i>there</i>
i son	di son	<i>here</i>
i-	di-	(passive prefix)
i allang	diallang	<i>eaten</i>
i allangna	ni allangna	<i>eaten by him</i>
ibahen	dibahen	<i>done</i>
i boan	diboan	<i>brought</i>
i buat	dibuat	<i>taken</i>
i dok	didok	<i>said; mentioned</i>
i pangkulingi	dipangkulingi	<i>spoken to</i>
i pingkir	dipingkir	<i>thought</i>
i lean	dilean	<i>given</i>
i tingkir	ditingkir	<i>peeped at; seen</i>
imana	ibana	<i>he/she</i>
inna	ninna	<i>he/she said; he/she did thus</i>
intor	pintor	<i>immediately</i>
jagul	jagung	<i>maize</i>
jai	jadi	<i>so; therefore</i>
jumpa	dapot	<i>found</i>
kaling	kaleng	<i>tin; measure of about 16kg of rice</i>
kansang*	hansang	<i>peanut</i>
koro	dope	<i>yet</i>
ndang koro	ndang dope	<i>not yet</i>
lak	laos	<i>just; simply</i>
lok	loas	<i>let it be</i>
mabe	masak	<i>ripe</i>
mana	nama	<i>only</i>
meser	meret; morot	<i>shifted</i>
musengani	muse	<i>then; later on</i>
nai	nari	<i>still; more</i>
ndara	ndada	<i>not</i>
nei	ndi	<i>this one here</i>
ni	ndi	<i>this one here</i>
ni	ni i	<i>of it</i>

Samosir Dialect	Standard Toba-Batak	English
nion	ndi on	<i>this one</i>
non	ndi on	<i>this one</i>
nungaeng	nuaeng	<i>now</i>
nungon ni muse	dungkon ni i muse	<i>and then...</i>
nung ke	dung pe	<i>after that; only then</i>
on	son	<i>here</i>
i on	di son	<i>here</i>
onde	ondeng	<i>the aforementioned</i>
poro	dope	<i>yet</i>
ndang poro	ndang dope	<i>not yet</i>
songonari*	saonnari	<i>now</i>
suntol	sungkot	<i>stuck (of objects)</i>
tea; manea	antan; mangantan	<i>to try the weight of</i>
tiak	teor	<i>to urinate</i>
tindang	jongjong	<i>to stand</i>
tinjang	jongjong	<i>to stand</i>
tonggor; manonggor	tondur; mainondur	<i>to watch (a performance)</i>
unga	nunga	<i>already</i>
uruk	dolok	<i>hill; high ground</i>

REFERENCES

ARITONANG, ROSPI

- 1981 Struktur fonem dalam bahasa Batak Toba. Skripsi Sarjana, Jurusan Bahasa Indonesia, Fakultas Sastra, Universitas Sumatera Utara, Medan.

BIRO PUSAT STATISTIK

- 1981a *Penduduk Sumatera menurut Propinsi dan Kabupaten/Kotamadya. Seri: L no 4. Hasil pencacahan lengkap sensus penduduk 1980.* Jakarta.
- 1981b *Peta indeks Kabupaten/Kotamadya dan Kecamatan di Indonesia. Seri: P no 1. Hasil pemetaan sensus penduduk 1980.* Jakarta.

BIRO PUSAT STATISTIK, KANTOR STATISTIK PROPINSI SUMATERA UTARA

- 1981 *Penduduk Propinsi Sumatera Utara 1980 - hasil pencacahan lengkap.* Medan.

BRAASEM, W.A.

- 1951 *Proza en poezie om het Heilige Meer der Bataks - De Bataks en hun volksletterkunde.* Djakarta, Amsterdam, Surabaya: De Moderne Boekhandel Indonesia.

COOL, F.G.

- 1921 Samosir. *Indië* 4/48, 2 Maart 1921:756-761; 4/49, 9 Maart 1921:773-778. Leiden.

CUNNINGHAM, Clark E.

- 1958 *The postwar migration of the Toba-Bataks to East Sumatra*. New Haven, Conn.: Southeast Asia Studies, Yale University.

DIJK, P.A.L.E. van

- 1896 Een tochtje per prauw langs Zuid-Samosir van Nainggolan tot Lottung, en een korte beschrijving van den oostelijke oever van het Tobameer van Siregar tot Poerba. *Tijdschrift van het Koninklijk Nederlandsch Aardrijkskundig Genootschap (TKNAG)*, 2e serie, 13:419-436. Leiden.

HARIARA, J.M.

- n.d. *Bunbunan ni hata Batak na marlapatan*. Pematang Siantar: Perpustakaan Theologia. Typescript, Universitas HKBP Nommensen.

HOEKSTRA, J.F.

- 1896 Onze kennis van het schiereiland Samosir en van het Tobameer in 1895. *TKNAG* 2e serie, 13:73-80. Leiden.

HOETAGALOENG, W.M.

- 1926 *Poestaha taringot toe Tarombo ni Bangso Batak*. Lagoeboti: Zendings-Drukkerij.

HURIA KRISTEN BATAK PROTESTAN

- 1981 *Notulen Sinode Godang H.K.B.P. ke-45, 31 Januari s/d 7 Februari 1981 di Seminarium Sipoholon, Tarutung*. Pearaja-Tarutung.

HUTABARAT, O.H.

- 1982 Pandidion di Distrik Samosir. "I pe asa tangkas huida". *Immanuel*, thn 92, 7/5:28-29. Pearaja-Tarutung.

HUTAJULU, V.S.M.

- 1982 Prospek Pulau Samosir dalam Pengembangan Industri Pariwisata di Wilayah Sumatera Utara. *Immanuel*, thn 92, 9/6:29-32. Pearaja-Tarutung.

HUTASOIT, M.

- 1978 *Ruhut Panurathon di Hata Batak dohot Turiturian*. Tarutung.

JOUSTRA, M.

- 1910 *Batakspiegel*. Uitgaven van het Bataksch Instituut no.3. Leiden: S.C. Van Doesburgh.

NABABAN, P.W.J.

- 1981 *A grammar of Toba Batak*. Materials in Languages of Indonesia 6. PL, D-37.

PARKIN, H.

- 1978 *Batak fruit of Hindu thought*. Madras: The Christian Literature Society.

PARLINDUNGAN, MANGARADJA ONGGANG

- 1964 *Pongkinangolngolan Sinambela gelar Tuanku Rao - Terror Agama Islam Mazhab Hambali di Tanah Batak. 1816-1833*. Djakarta: Penerbit Tandjung Pengharapan.

PERCIVAL, W.K.

- 1981 *A grammar of the urbanised Toba-Batak of Medan.* PL, B-76.

PLEYTE, C.M.

- 1895 *De verkenning der Bataklanden - Eene bijdrage tot de geschiedenis der ontdekking van het Toba-meer.* TKNAG, 2e serie, 12:71-96; 727-739. Leiden.

SARUMPAET, J.P.

- 1982 *Linguistic varieties in Toba-Batak.* In Amran Halim, Lois Carrington and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, vol.3: *Accent on variety*, 27-78. PL, C-76.

SCHREIBER, A.

- 1874 *Die Battas in ihrem Verhältnis zu den Malaien von Sumatra.* Inaugural-Dissertation, Universität Leipzig. Barmen: J.F. Steinhaus.

SIAHAAN, A.B.

- 1982 *Barita Zending.* *Immanuel*, thn 92/6, April, 25-30. Pearaja-Tarutung.

SINAGA, ANICETIUS B.

- 1981 *The Toba-Batak High God - transcendence and immanence.* *Studia Instituti Anthropos*, vol.38. St Augustin, West Germany: Anthropos Institute.

SIREGAR, AHMAD SAMIN

- 1977 *Kamus bahasa Angkola/Mandailing - Indonesia.* Jakarta: Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan.

TUUK, H.N. van der

- 1864 *Tobasche Spraakkunst, voor het Nederlandsch Bijbelgenootschap. Eerste Stuk (Klankenstelsel).* Amsterdam: Frederik Muller.

VOLZ, W.

- 1899 *Zum Toba-See in Central-Sumatra.* TKNAG, 2e serie, 16:415-485. Leiden.

VOORHOEVE, P.

- 1955 *Critical survey of studies on the languages of Sumatra.* The Hague: Martinus Nijhoff.

WARNECK, J.

- 1977 *Toba-Bataksch-Deutsches Wörterbuch.* The Hague: Martinus Nijhoff.

WURM, S.A. and Shirô HATTORI, eds

- 1983 *Language atlas of the Pacific area, part 2: Japan area, Taiwan, Philippines, mainland and insular south-east Asia.* Canberra: The Australian Academy of the Humanities, in collaboration with the Japan Academy. PL, C-67.

THE BARRIER ISLAND LANGUAGES IN THE AUSTRONESIAN LANGUAGE FAMILY

Bernd Nothofer

1. QUANTITATIVE AND QUALITATIVE EVIDENCE AS BASIS FOR SUBGROUPING ARGUMENTS

Subgrouping arguments can be based on quantitative or on qualitative evidence.¹ Quantitative evidence consists of the statistical study of the vocabularies of languages. Qualitative evidence consists of the collection of exclusively shared innovations. As we will see below, some scholars appeal to both quantitative and qualitative evidence in determining subrelationships, giving preference to qualitative evidence whenever it conflicts with quantitative evidence. The fact that there exists a conflict between these two kinds of evidence shows that we have to question either the assumptions of lexicostatistics or of the comparative method. Blust (1981) irrefutably disproves one of the fundamental assumptions of current lexicostatistical theory, namely that basic vocabulary gets replaced at a rate which is constant for all languages at all times. Blust observes retention percentages from 58.5% to 15.8% in his sample of 55 languages and dialects. It therefore appears that only qualitative evidence represents a reliable basis for the determination of subrelationship.

2. AUSTRONESIAN SUBGROUPING AND THE POSITION OF THE BARRIER ISLAND LANGUAGES IN THE AUSTRONESIAN LANGUAGE FAMILY

Only few scholars who have dealt with the subgrouping of the Austronesian language family included the Barrier island languages in their study. The first one was Brandstetter who concluded that Nias was most closely related to Malagasy. This hypothesis was rejected by Lafeber (1922:57-58) who also recognised "strange phonetic agreements" between Malagasy and Nias "which also appear in other Barrier islands such as the occurrence of the sequence ndr (as reflex of *nD or *nd - BN), of f (as a dialect of Enggano) as reflex of *p and of h (as in Enggano, Toba and Mandailing) as reflex of *k". Lafeber argued that "the Malagasy vocabulary is much closer to the Malay lexicon than to that of Nias". He claimed that the vocabulary of "Batak-Gayo" has many agreements with that of "Nias - Simalur - Mentawai - Enggano". Unfortunately, he gave only two examples:² TBt. sada, Ga. södö, sara, Ni. sara, Me. sara, Sim. sara *one*; TBt. toru *below* and its cognates in Gayo and the Barrier islands. However, Lafeber never fulfilled his promise to present further lexical evidence for his hypothesis, since the announced second volume of his book in which this evidence was to be given never appeared in print.

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 87-109. *Pacific Linguistics*, C-94, 1986.

© Bernd Nothofer

In his analysis of Mentawai Adriani (1928) reached the conclusion that "one cannot say that the language of the Mentawai islands does not seem to be in its place in its environment. Mentawai is a language which - regarding its geographical position - has no strange character".

Adriani quoted Jonker (1918) who wrote an article on Mentawai for the *Encyclopaedie van Nederlandsch Oost-Indië*:

Mentawai is certainly related to Nias, but it is notably different, which is due in part to the fact that its sound system has been retained more completely; the difference in the lexicon is very big. Generally, Mentawai words make a strange impression; many items of general vocabulary must have been lost and replaced by others.

In the *Atlas van tropisch Nederland* Esser (1938) presented a single-page classification of the languages of the then Netherlands East Indies. He recognised 17 groups of AN languages. One of these is the "Sumatra" group which consists of Aceh, Gayo, Batak dialects, Minangkabau, (Lubu), Malay, Middle Malay, Rejang-Lebong, Lampung, Simalur, Nias, (Sichule), Mentawai, Enggano, Loncong, Lom, Orang Laut.

Neither in his *Grammatischer Abriß des Enggano* (1940) nor in his *Untersuchungen über die Laut-, Wort- und Satzlehre des Nias* (1937) did Kähler comment on the relationships of these languages to other AN languages. However, in the introduction to his unpublished *Grammatik der Simalursprache* (n.d., probably written in the late 1930s), Kähler wrote that "the Simalur vocabulary contains such a clearly recognisable Celebes-Philippine substratum that a formerly close contact between Simalur and this northern language group is certain". In the fifth section of his manuscript which is entitled "Borrowings in Simalur and their implications" Kähler lists what he treats as loans from 1) Gayo, Aceh; 2) Minangkabau, Batak dialects; 3) Sundanese, Javanese; 4) Celebes and Philippine languages; 5) Borneo languages; 6) languages in the east of the archipelago. Since the largest body of evidence was accumulated for the Celebes and Philippine languages, Kähler drew the conclusion that

... a formerly close connection between the inhabitants of these areas seems certain. This common vocabulary cannot be treated as single borrowings, since they consist partly of the oddest words. Simalur shares the possession of a linguistic substratum originating from the northern language group of Indonesia with other dialects on the islands on the west coast of Sumatra (Sichule, Nias, Mentawai, Enggano), although Nias has more words and Mentawai mostly different words which originate from the Celebes group. This original substratum in the lexicon of Simalur was later superimposed by a Sumatran layer. ... In my opinion, the settlement of Simalur (and of the other Barrier islands) cannot have taken place via Sumatra, because those words which appear in the island languages and which originate from this northern group (Celebes-Philippines) do not exist in dialects of Sumatra, although some of them have a lexicographically mixed character.

Before commenting on Kähler's hypothesis, I would like to describe a work which was written by Willms (1955), a student of Kähler. In his analysis of Mentawai Willms compiled lists of what he treated as borrowings from languages

of Celebes and Sumatra. He furthermore has a list of so-called Mentawai borrowings from Nias and Simalur.

Both Kähler and Willms automatically treated all those Simalur and Mentawai words that do not reflect a PAN etymon reconstructed by Dempwolff (1934-1938) as borrowings from one of the languages in which a related form occurs. For example, Willms reconstructed a form *əRəm in order to account for KBT. ərəm, TBT., Angk. orom, Me. om *to resist*. Instead of treating Me. om as a cognate, which actually points to a reconstruction with *R (and not *r), he treated it as a borrowing from the Batak languages and argued that "Mentawai had contact with Batak before the sound change *R > Batak languages r occurred".

Neither Kähler nor Willms considers the possibility that the words which Barrier island languages seem to share exclusively with each other, with Sumatran or with Sulawesi-Philippine languages might reflect an etymon of their respective last common proto-language. There can be no doubt that particularly in the case of words which seem to be shared exclusively by a Barrier island language and neighbouring Barrier island languages or by a Barrier island language and Sumatran languages there in fact exists a borrowing relationship. This is a more difficult argument in the case of the many words listed by Kähler and Willms which appear to be shared exclusively by Barrier island and Sulawesi-Philippine languages.

If we interpret these as reflections of etyma of an earlier common proto-language which is not PAN, one might indeed argue that these two language groups have an exclusively shared history. Although a close examination of the lists compiled by the two German scholars shows that in a considerable number of cases either the forms or the meanings are too different to allow a treatment as cognates or there exist cognates in non-Barrier island and non-Sulawesi-Philippine languages, there remain some interesting comparisons which could be treated as lending support to such an argument.

Salzner (1960) who wrote the *Sprachenatlas des Indopazifischen Raumes* included the Barrier island languages in his so-called "Sumatra group" of south-west Indonesian languages. This group is almost identical with that of Esser (1938). It contains Aceh, Gayo, Batak languages, Minangkabau, Malay, Rejang-Lebung, Middle Malay, Lampung, Lom = Mapor, Basa Loncong, Simalur, Nias, Mentawai, Enggano, and Samsam.

In 1965 Dyen published his *A lexicostatistical classification of the Austronesian languages*. In this study the Austronesian language family is divided into 40 first-order subgroups. Most of them are located in western Melanesia and adjacent areas. We also find one in northern Formosa and another one on Enggano. Blust (1981:13) commented on these results as follows:

The existence of lexicostatistically-defined first-order subgroups in more than one widely separated area must - if the percentages accurately reflect the historical order of splits - be explained on a hypothesis of migration. Given Dyen's methodological assumptions and the reported percentages it would appear simplest to explain the location of the Atalyic Subfamily and Enggano as a result of several migrations from western Melanesia which resulted in long-distance settlements to the north and west. However, Dyen did not adopt such a hypothesis. Instead, in the case of Enggano he attempted (p.56) to find intermediate percentages that link this language with other languages of western Indonesia. An examination of lists for Enggano's northern neighbours

Mentawai and Nias (neither of which was considered in the classification proper) failed to provide such intermediate percentages. Dyen admits that the explanation for the low cognate percentages connecting Enggano, Mentawai and Nias with each other and with other AN languages is not clear. Nonetheless he believes "... it is likely that these languages will ultimately prove to be closely related to the languages of western Indonesia by a non-lexicostatistical argument. This is suggested by the appearance of Mentawai buluk, Nias bulu *leaf* corresponding to Toba Batak bulu *leaf* (cf. the almost universal cognates of Tagalog da:hon *leaf*), Mentawai ka-baga, Nias bacha *in* corresponding to Toba Batak di-bagas-in *in*, Mentawai urat *root* corresponding to Toba Batak urat *root* (cf. the widespread cognates meaning *vein, tendon*), and of Nias f-al-ea *lie down* corresponding to Toba Batak peak *lie down*."

In footnote 8 of this article Blust demonstrated that Dyen's qualitative evidence does not always hold. Cognates of the forms for *leaf* are widespread in the Philippine languages and a reconstruction *bulu *foliage* had already been proposed by Dempwolff. Similarly, forms which continue *uRat *vein, tendon* in the meaning *root* occur not only in Mentawai and Toba-Batak but also in many Borneo languages (e.g. Maloh urat *vein, root*).

Furthermore, cognates of the Mentawai, Nias and Toba-Batak forms for *in* also occur in Philippine languages (e.g. Tag. sa-balas (*inland* =) *north-west*) and also in this case a reconstruction was in fact proposed by Dempwolff (*bajas *interior*).

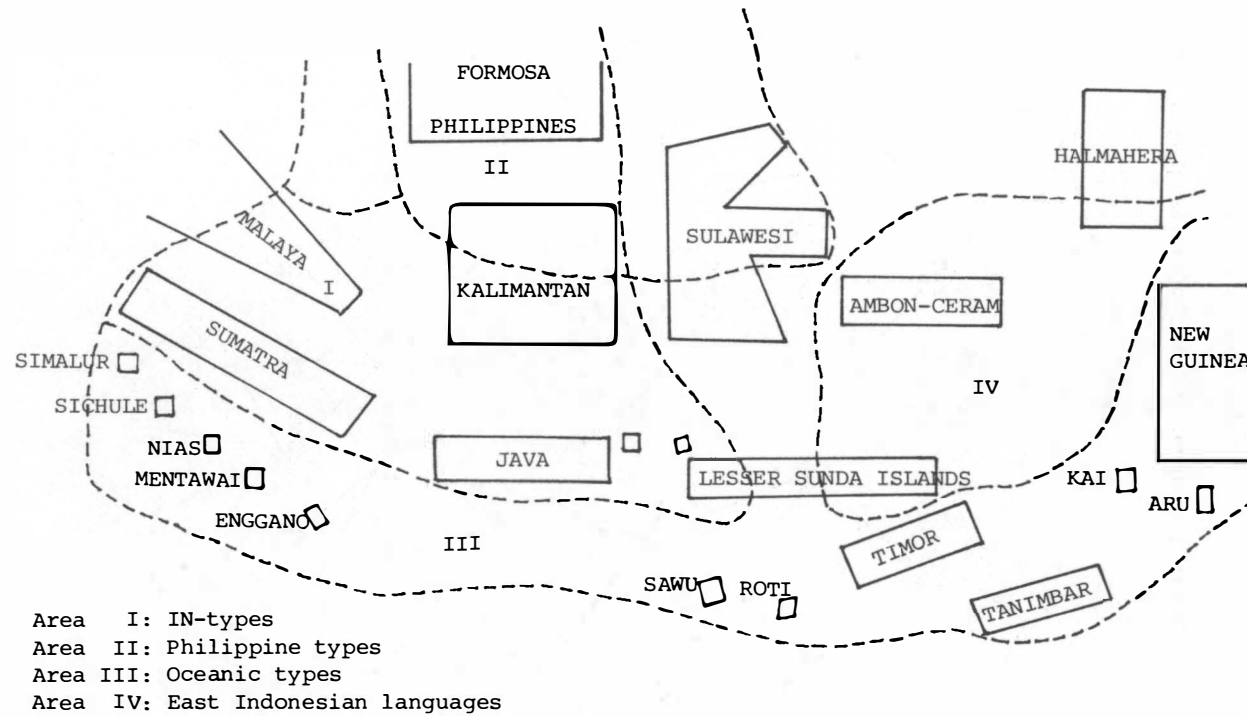
Blust did not attempt to subgroup the Barrier island languages although he wrote in the footnote cited above: "Although I am entirely in sympathy with Dyen's attempts to link Enggano, Mentawai and Nias with other languages of western Indonesia ...".

Capell (1982) argued that

Enggano is not an Austronesian language from the point of view of its vocabulary and its grammar ... Enggano is structurally *sui generis*; ... it does not have Melanesian traits as for example Mentawai ... Enggano is a remnant of these pre-IN languages, which indeed has IN borrowings, but remains non-Austronesian.

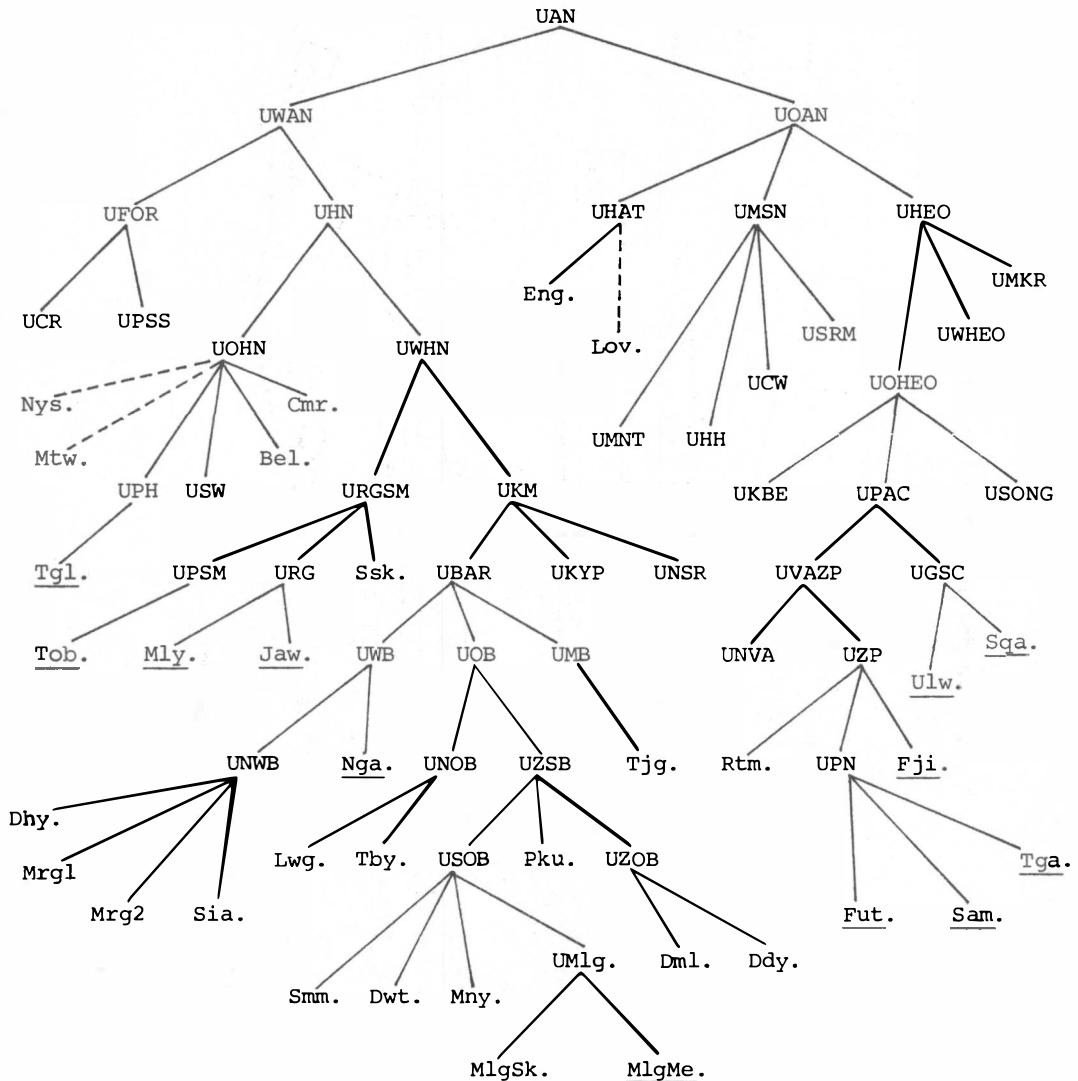
Finally, Capell arrived at a distinction of four language-types in Indonesia. The arguments for these distinctions and for the grouping of the Barrier island languages as being members of the Oceanic type remain unclear to me. Capell drew the following diagrammatic map (1982:15):

A location diagram of the Indonesian area



The most recent attempt at a subgrouping of the Barrier island languages is Mahdi's manuscript "Morphophonologische Besonderheiten und historische Phonologie des Malagasy" which I received in April 1984. Mahdi divides the AN languages into two primary groups: 1) Proto-West-Austronesian and 2) Proto-East-Austronesian. Nias and Mentawai belong to 1) and Enggano to 2) (see Mahdi's tree-configuration).

Mahdi's subgrouping of the AN languages



List of language and dialect abbreviations used in Mahdi's diagram (language names in English, here):

Bel	Belau (Palau)	GSC	Nggela-San Cristobal
Cmr	Chamorro	HAT	Hartanic
Ddy	Dusun-Dejah	HEO	Heonesic
Dhy	Dohoi	HH	Halmaheran
Dml	Dusun-Malang	HN	Hesperonesian
Dwt	Dusun-Witu	KBE	Kimbe
Eng	Enggano	KM	Kalimantan
Fji	Fijian	KYP	Kayan-Punan
Fut	Futuna	MB	Mahakam-Barito
Jaw	Javanese	MKR	Micronesian (Kern)
Lov	Lovaia	MNT	Molucco-Nusatenggara
Lwg	Lawangan	MSN	Mesonesian
MlgMe	Mérina (Malagasy)	NOB	North-east Barito
MlgSk	Sakala'va (Malagasy)	NSR	North Sarawakian
Mly	Malay	NVA	North Vanuatuan
Mny	Ma'anjan	NWB	North-west Barito
Mrg1	Murung 1 (Hudson 1967)	OAN	East Austronesian
Mrg2	Murung 2 (Hudson 1967)	OB	East Barito
Mtw	Mentawai	OHEO	East Heonesic
Nga	Ngadju	OHN	East Hesperonesian
Nys	Nias	PAC	Pacific
Pku	Paku	PH	Philippines
Rtm	Rotuma	PN	Polynesian
Sam	Samoan	PSM	Paleo-Sumatran
Sia	Siang	PSS	Paiwano-Saisiat
Smm	Samihim	RG	Urangic
Sqa	Sa'a	RGSM	Urango-Sumatran
Ssk	Sasak	SOB	South-east Barito
Tby	Taboyan	SONG	South-east New Guinean
Tga	Tongan	SRM	Sarmic
Tgl	Tagalog	SW	Sulawesic
Tjg	Tundjung	U	Ur- (= Proto-)
Tob	Toba-Batak	VAZP	Vanuatan-Central Pacific
Ulw	Ulawa	WAN	West Austronesian
AA	Austroasiatic (non-AN)	WB	West Barito
AN	Austronesian	WHEO	West Heonesic
BAR	Barito	WHN	West Hesperonesian
CR	Tsou-Rukai	ZOB	Central East Barito
CW	Cenderawasih (Geelvink)	ZP	Central Pacific
FOR	Formosan	ZSB	Central South Barito

The family tree shows that Nias and Mentawai directly continue Proto-East Hesperonesian just as do Proto-Philippine, Proto-Sulawesi, Palau and Chamorro. Proto-East Hesperonesian and Proto-West Hesperonesian directly continue Proto-Hesperonesian which together with Proto-Formosan is a daughter language of Proto-West Austronesian. Enggano and Lovaia (East Timor) are grouped as daughter languages of Proto-Hartanic which in turn directly continues Proto-East Austronesian.

Mahdi (n.d.:58) comments on his subgrouping by writing that

... the Philippines and parts of west and central Indonesia were inhabited by peoples speaking East Austronesian languages. Because they were superseded by West Austronesian

languages most of their languages were either lost or are preserved only as substratum, e.g. in the languages of the islands off the coast of west Sumatra, in the Batak dialects of Sumatra, in the Aeta dialects and some other idioms of the Philippines, Sulawesi and Nusa Tenggara. It is significant that these idioms often have reflexes of *qa(R)[C]a as the word for *man* For the time being, I will assume that the languages which were here lost form a separate subdivision of the East Austronesian group, the proto-language of which I will call Proto-Hartanic. It is indeed possible that Enggano might be regarded as a direct daughter language of Proto-Hartanic. The same possibly also holds for Lovaia.

To comment on Mahdi's last point first: he probably considers Enggano and Lovaia as belonging to the same subgroup, because in both languages *t, *C > k and *s > t.

Mahdi's subgrouping seems to agree partly with the hypotheses put forth by Kähler, Willms and maybe Capell. I assume that the grouping of the Barrier island languages with the Sulawesi-Philippine and/or the Oceanic languages is based on the observation that there exists a number of etyma which have cognates only in these languages. However, this observation is only of relevance for subgrouping, if the etyma whose cognates have this distribution are *innovations*. There is, however, no good reason to believe that e.g. *qa(R)[C]a has replaced a form that represented the same meaning in PAN.

3. COMPETING VIEWS ON THE HISTORY OF MENTAWAI CULTURE

The Swiss anthropologist Schefold who wrote various articles on the religion of Mentawai (1972, 1976) maintained in his book *Speelgoed voor de zielen* (1979: 13) that

... according to anthropological and linguistic studies the people of Mentawai are closely related to the non-Islamised tribes (the Batak) on Sumatra. This supports the hypothesis that the first Mentawai people came from Sumatra. The time of this arrival can only be given approximately. The people of Mentawai do not know how to work metal, they have no knowledge of rice-planting or weaving. Their culture must therefore be older than the bronze age.

In another article (1979:201) Schefold claimed that "metal working and rice-planting came to west and central Indonesia at the same time, but after a neolithic Austronesian migration which also influenced eastern Indonesia". Furthermore, Schefold (1979:13) argued that

... there are also elements lacking in Mentawai which one can ascribe to the late neolithicum on the basis of the situation in Polynesia: the society is egalitarian, there are no chiefs; the Mentawai people do not know the erection of megaliths. The Mentawai islands represent an early tradition in the neolithicum.

It is interesting to note that Marschall (1966) regarded the Mentawai culture as recessive which secondarily gave up metal-working, rice-planting and weaving.³ Marschall's hypothesis supports Blust's reconstructions of PAN etyma for metal, rice and weaving.

4. QUALITATIVE EVIDENCE FOR A BARRIER ISLAND-BATAK SUBGROUP

In the following pages I will show 1) that strong qualitative evidence can be adduced in support of a Barrier island-Batak subgroup and 2) that this subgroup contains all Barrier island languages, perhaps including Enggano. Because of the lack of data it is difficult to provide substantial evidence for grouping Enggano with these languages.

The evidence will consist of exclusively shared phonological and lexical innovations. Exclusively shared phonological innovations are insufficient for the establishment of a subgroup, since the number of possible sound changes is rather limited compared to the number of possible lexical changes. It follows that identical sound changes which occur in geographically distant languages or language clusters cannot be taken alone as evidence for an exclusively shared history of these languages. It is for this reason that e.g. the occurrence of g as reflex of *j in two geographically distant language groups such as the Barrier island-Batak group and the Philippine group is interpreted as two separate innovations for the time being. Further evidence, be it grammatical, lexical or semantic, has to be adduced. If we based our analysis on phonological innovations alone, Enggano would probably be subgrouped with a language such as Douru (spoken in the Central District of Papua): *t > En., Dou. k; *k > En., Dou. Ø; *s > En., Dou. t; *ŋ > En. h, Dou. Ø.

The material for the island languages consists mostly of grammars and dictionaries written by Kähler (1937, 1940, 1959, 1961, 1975). Other important information appears in Morris 1900 and Zainuddin HR Lenggang 1978 for Mentawai and Sundermann 1905 for Nias. None of these works contains reliable material on the phonology of the languages examined. Toba-Batak material is taken from van der Tuuk 1971 and Warneck 1906. During two fieldtrips to Mentawai I collected Swadesh lists for Mentawai dialects. For Nias I was sent Swadesh lists of six dialects by German missionaries. These lists were used in a lexicostatistical calculation of the cognate percentages among Mentawai and Nias dialects respectively. The results for Mentawai are listed in Table 1:

Table 1: Lexicostatistical percentages among the Mentawai dialects

	Simatalu	Terekan	Sikabaluan	Saxaliow	Sikakap	Sipora
Simalegi	69	74	70	57	62	58
Simatalu		71	71	65	62	61
Terekan			71	58	57	57
Sikabaluan				60	61	60
Saxaliow					56	61
Sikakap						95

There is relatively little dialect variation among the dialects of Nias (cognate percentage about 80%). As we can see from Table 1 this also holds for the dialects of the southern Mentawai islands. The dialects of Siberut however are very different from each other (cognate percentages varying between 71% and 57%) and from the south Mentawai dialects (cognate percentages varying between 62% and 57%).

4.1 The phoneme inventories of the languages under investigation

The most difficult part of the phonemic analysis of the island languages is Kähler's treatment of the vowels represented by the symbols ə, ö, õ, u, ē. It appears that ö, õ, u, ē are phonetically [ɨ]. Another problematic symbol is Kähler's ʰ which appears to be [ç]. In languages which also have [x], [ç] and [x] seem to be in complementary distribution.

4.1.1 The Simalur phoneme inventory

Simalur has the following seven vowel phonemes according to Kähler:

i	ɨ	u
e	ə	o
	a	

Nasal vowels are in free variation with their corresponding oral vowels. They only occur very rarely and only in the environment of nasal consonants.

Simalur has the following consonant phonemes:

p	t	c	k	ʔ
b	d	j	g	
m	n	ɲ	ŋ	
f	s		x	h
	l			
	r			
w	y			

The phoneme /x/ has the allophones [x] and [ç]. The latter occurs in the environment of /i/, /ɨ/ or /e/.

4.1.2 The Sichule phoneme inventory

The vowel phonemes of Sichule are, according to Kähler:

i		ɨ	u
e	ö	ə	o
		a	

The consonant phonemes are:

p	t	k	
b	d	g	
m	n	ŋ	
f	s	x	h
	l		
	r		
	y		

Again, the phoneme /x/ has the allophones [x] and [ç]. /b d g/ in final position are realised as unreleased stops.

4.1.3 The Nias phoneme inventory

Nias has six vowel phonemes:

i	ɨ	u
e		o
	a	

Its consonant phonemes are:

	t	k	ʔ
b	d	g	
m	n	ŋ	
f	s	x	h
v	z		
	l		
	r		
w	y		

4.1.4 The Mentawai phoneme inventory

Mentawai has the following five vowel phonemes:

i	u
e	o
	a

The consonant phonemes are:

p	t	c	k	ʔ
b	d	j	g	
m	n	ñ	ŋ	
	s			
	l			
	r			
w		y		

4.1.5 The Enggano phoneme inventory

The vowel phonemes of Enggano are:

i	ɨ	u
e		o
ɛ	ɔ	
	a	

According to Kähler each oral vowel phoneme has a corresponding nasal vowel phoneme:

ĩ	ɨ̃	ũ
ẽ		õ
ɛ̃	ɔ̃	
	ã	

In his Simalur and Sichule dictionaries Kähler does not distinguish between e and ε or between o and ɔ.

The number of consonant phonemes depends on the dialect:

p	(t)	c	k	ʔ
b	d	(j)		
m	n	n		
(f)				h

The phonemes in parentheses only appear in the southern dialects.

4.1.6 The Toba-Batak phoneme inventory

The vowel phonemes are:

i	u
e	o
a	

Toba-Batak has 14 consonant phonemes:

p	t		k
b	d	j	g
m	n		n
	s		h
	l		
	r		

4.2 Phonological history of the languages under investigation

We will not give a full account of the phonological history of each of the languages from reconstructed material. Instead we will present a table which consists of a general overview of the PAN phonemes and their reflexes in the six languages (Table 2).

4.3 Phonological innovations and irregularities shared among the six languages

In this section we will deal 1) with the phonological innovations and 2) with the phonological irregularities which are shared among Simalur, Sichule, Nias, Mentawai, Enggano and Toba-Batak.

Table 2: Phonological changes

PAN a i u ə	Mentawai a i,e u,o e,o	Nias a,o i,e u,o t,o	Toba-Batak a i,e u,o o	Simalur a i u ə,t	Enggano {ə} {ɨ} {ʉ} {ɔ'}{ɔ'}	Sichule a i,e u,o ə,ɔ',t
b	b-/m -b- -p	b- -v- -θ	b- -b- -p	f- -f- -(b)	p- -p- -θ	b-/f- -f- -θ
d	r- -r- -t	d- -r- -θ	d- -d- -t	d-/r- -d-/r- -(?)	d- -d- -θ	d- -(?)- -(?)
θ	r- -r- -t	d-/r- -r- -θ	d- -d- -r	d-/r- -r- -r	d-/r- -d- -θ	d-/l- -l- -n(?)
Z- -Z-	r	r	d	d- -r- -r	r- -d- -θ	d- -l- -l
z- -z-	j	z	j	j- -j- -j	?	?
g	g- -(?)- -(?)	g- -g- -θ	g- -j- -k	g- -g- -(?)	?	g- -g- -θ
-j- -j	-g- -u	-x- -θ	-g- -k	-h-/-x- -(g)/?	-ç-/-θ-/-h- -θ	-x- -θ
R	θ	θ	r	l/θ	θ	θ
h	θ	h-/θ -h-/θ- -θ	θ	θ	θ	θ
q	θ	h-/θ -h-/θ- -θ	θ	θ- -θ- -(h)	h- -θ- -θ	θ- -h-/θ- -θ
y- -y- -ay	(?) -j- -e	y- -y- -e	(?) -θ- -e	(?) -e- -ae	(?) - - (?) - -ae	(?) - -e-/-y-? -ae
-əy	-ey	-e	-e	-ae	- (?)	-e
-uy	-uy/-i	-i	-i	-oe	- (?)	-i
k	k- -k- -?	?-/θ- -?-	h- -h- -k	θ-/k- -?-/-k-/-θ - (g)	?-/k-/-θ- -θ- -θ	?- -?- -θ
c- -c-	s	s- -(?) -	s	s	?	?
l	l- -l- -lV	l- -l- -θ	l	l	l-/r-/d- -l-/r-/-d- -θ	l- -l- -n(?)
r	r- -r- -rV	r- -r- -θ	r	r	(?) - -d- -θ	l- -l- - (?)
m	m- -m- -m /-p	m- -m- -θ	m	m	b- -m-/-b- -θ	m- -m- -θ
n	n- -n- -n /-t	n- -n- -θ	n	n	(?) - -d-/-r- -θ	n- -n- -θ
ñ- -ñ-	(?) - -n-	n	n	n	-n-	(?) - -n-
ŋ	ŋ- -ŋ- -ŋ /-k	ŋ- -ŋ- -θ	ŋ	(?) - -ŋ- -ŋ	(?) - -h-/-θ- - (?)	ŋ- -ŋ- -θ
p	p- -p- -m /-p	f- -f- -θ	p	θ- -h-/-x-/-θ- -(b)	p- -p-/-b- -θ	f- -f- -θ
t	t- -t- -n /-t	t- -t- -θ	t	t- -t- -(d)	k- -k- -θ	t- -t- -θ
T- -T-	t	t	t	t-	?	?
s	s- -s- -θ	s- -s- -θ	s	s	k- -k- -θ	θ- -h-/-θ- -θ
w- -w- -aw	b- -b- -au -ou -eu -eu	w- -w- -o/-ö	θ- -θ- -o	(?) - -w- -ao	b- -b- - (?)	θ- -w- -i,ao (?)
-əw	-eu	-ö	-o	- (?)	- (?)	- (?)

4.3.1 Shared phonological innovations

VOWEL AND DIPHTHONG SHIFTS

PAN	SIMALUR	SICHULE	NIAS	MENTAWAI	ENGGANO	TOBA-BATAK	
1. *e	ə, ɨ	ə, ɔ̃, ɨ	i		ɔ̃	o	
Examples:							
*telu	təlu, təlo	tölu	tɨlu	(telu)	ʔakoru	tolu	<i>three</i>
*(b)eli	bəli	bɨli	bɨli=ɨli		e-odi <i>price</i>	boli <i>bride</i>	<i>buy</i>
2. *e after *R	e				ɛ̃		
Example:							
*Sa-ReZan	aeran	(ola)	(ora)	(orat)	e-hěă	(ardan)	<i>ladder, staircase</i>
3. *e before *j	o			o			
Example:							
*qunej	unog	(unɨ)	(hunɨ)	unou		(unok)	<i>marrow</i>
4. *-ay	ae	ae			ae		
Examples:							
*kuday	kudae		(kude, gude-gude)	(ore)	e-ʔorae		<i>basket made of bamboo</i>
*baday	badae	badae	(bade)				<i>storm</i>

VOWEL MERGERS

1. *au		ɨu	ɨu	eu			
*eu		ɨu	ɨu	eu			
Examples:							
*Zauq	(dao)	a-dɨu	a-rɨu	a-reu		(dao)	<i>far</i>
*behew	(fo)	bɨu	bɨu	beu	(upau)	(bau)	<i>smell</i>
2. *a before *-k and *-ŋ		o	o				
*e before *-k and *-ŋ		o	o				
Examples:							
*anak		n-ono	n-ono		(e-ara)	(anak) son	<i>child</i>
*ma(n)yaŋ	(maeaŋ)	mõẽõ	moyo	(mañan)			<i>hawk</i>
*qutek	(utaʔ)	uto	uto	(ute) head		(utok-utok)	<i>brain, marrow</i>
*lateŋ	(lalateŋ)	lato	lato	(lalatek)		latoŋ	<i>stinging nettle</i>
3. *a in the environment of o (> *a)		o	o				
*e in the environment of o (> *e)		o	o				

PAN	SIMALUR	SICHULE	NIAS	MENTAWAI	ENGGANO	TOBA-BATAK
Examples:						
*layaR	(laeal)	loyo	loyo	lajo		rear
*anak		n-ono	n-ono			<i>sail</i>
*DeŋeR		loŋo	roŋo			<i>child</i>
						<i>hear</i>
4. *a in the environment of *R						
*e in the environment of *R		o	o	o		
contraction of o's (> *a or *e) after the loss of *R						
Examples:						
*paRi	(ali)	foi	foi	poi		<i>stingray</i>
*qabaRaH	(bala)	bo	bo	bo		<i>shoulder</i>
*Ratus	(latus)	otu	otu	otu		<i>hundred</i>
*DeŋeR		loŋo	roŋo			<i>hear</i>
*timbeR	(tebəl)		simbo	timbo	(e-ipo)	<i>smoke</i>
*Sa-ReZan	(aeran)	ola	ora	orat	(e-hea)	(ardan)
*eRem				om		orom
*baRaH	(bala,fala)	bo (naiti)	bo(galiti)	bo		<i>resist</i>
						<i>heat, red</i>
CONSONANT SHIFTS						
1. *-j-	x,h,∅	x	x	g x(Pagai)	ç,h,∅	g
Examples:						
*Sua(n)ji	axi,ahi	axi	axi	bagi	(āhāi)	an̄gi
*sijem	sixəm	ixōm	sixi	sigep	e-kiço	<i>younger brother</i>
						<i>ant</i>
2. *-ñ-	n	n	n	n	n	n
Example:						
*peñu	əno	fɪnu	fɪnu	penu-ŋ	eʔünũʔünũ	ponu
						<i>sea-turtle</i>
3. *-ç		∅ (?)	∅		∅	
Example:						
*xe(m)pat	(ad)	ifa	ifa	(epat)	ʔa-opa	(opat)
						<i>four</i>
4. *-b-	f	f				
Example:						
*tabeq	tafi(x)	tafi	(tavɪ)	(tabe)		(tabo)
						<i>fat</i>

PAN	SIMALUR	SICHULE	NIAS	MENTAWAI	ENGGANO	TOBA-BATAK
5. *p- -p- Examples: *pitu *lapaR	(itu) (lahal)	f fitu olofo	f fitu lofo	(pitu)		(pitu) <i>seven hungry</i>
6. *-n[dD]- Examples: *lan[dD]aw *tanduk	(tadu?)		ndr tandru	ndr landrou	(e-kadu?u)	(lando) tanduk <i>limb, length horn</i>
7. *R Examples: *layaR *Rumaq	(laea1) (luma)	∅ loyo	∅ loyo omo	∅ lajo uma	∅ e-uba	rear (ruma) <i>sail house</i>
CONSONANT MERGERS						
1. *ñ *n Examples: *peñu *bunuq	n n əno bunu, funu	n n fɪnu bunu	n n fɪnu bunu	n n penu-ŋ munu	(e?ũnũ?ũnũ) (pudu)	n n ponu bunu <i>sea-turtle kill</i>
2. *-Z- *-D- Examples: *Sa-ReZan *peDem	r r aeran ɪrɪ?	l l ola mɪlɪ	r r ora mɪrɪ	r r orat merem	? (e-hěã)	d d ardan podom <i>ladder, staircase sleep</i>
3. *c- *s- Examples: *cimcim *caremin *sawa	s s sincim səremen sawa	? (awa)	s s sirɪmi sawa	s s sipsip sawa	? (e-dio) (e-nānĩ)	s s sormin sa <i>ring mirror snake</i>
4. *Z- ... *l *D ... *l *d ... *l Examples: *zalan *Dilaq *daliq	(dalan, ralan) (dila) (dalig)	l...l l...l l...l lala lela lali	l...l l...l l...l lala lela lali	l...l l...l l...l lalan <i>vu1va</i> lila	(e-dio) (e-nānĩ)	(dalan) dila <i>path tongue root</i>

4.3.2 Shared phonological irregularities

Another kind of phonological material which can also be taken as evidence for subgrouping is shared irregularities in phonological development. The first list contains irregularities which appear to be exclusively shared by Barrier island languages and the second list those which appear to be exclusively shared by at least one Barrier island language and Toba-Batak. Sometimes forms from other Sumatran languages are cited.

List 1: Phonological irregularities shared by Barrier island languages

- *pulaw > Sim., Sich. ulao, Ni. hulo *island* (*q- instead of *p)
- *lanaw > Sim. ɣali, Sich. ɣali=nalao, Ni. ɣali-ɣali *fly* (metathesis)
- *betuŋ > Ni. motu *kind of wood*, Me. metuk *kind of bamboo* (*m instead of *b)
- *uRat > Sich. g-uŋo *vein, tendon*, Me. uŋat *vein, tendon, root* (unexpected ŋ)
- *Cuqelaŋ/CuqelaN > Sich., Ni. tɪla *bone* (*e instead of *o)
- *beli > Ni. ɪli *buy*, En. e-odi *price* (loss of *b)

List 2: Phonological irregularities shared by Barrier island languages and Toba-Batak (and/or other Sumatran languages)

- *lalej > *lanej > Ni. nalö (< *nalej), Me. si-linau (< *si-linaj < *silenaj), TBt. lanok, KBt. lanəŋ *fly*
- *bibɪR > Me. bibo, KBt. bibər *lip* (*e instead of *i)
- *Rejaŋ > Me. ogdag *wooden stick to work coconut*, TBt. ordaŋ *planting-stick*, Mal. rəjaŋ *break up with a crowbar* (metathesis of *Re)

4.4 Lexical innovations shared among the six languages under investigation

The lexical evidence for a group consisting of the Barrier island languages and Toba-Batak is divided into two lists. Again, the first list contains lexical items which appear to be exclusively shared by Barrier island languages and the second list contains those which appear to be exclusively shared by at least one Barrier island language and Toba-Batak. Sometimes items from other Sumatran languages are cited.

List 1: Lexical innovations shared by Barrier island languages

- Ni. la-lau *to braid, plait, twist*, Me. lai *to wrap, wind, tie*
- Ni. xiti harita *young green beans*, Me. gette *kind of keladi (taro)*
- Ni. havo, Me. abo *bunch of bananas*
- Ni. alito, Me. alito *fire*
- Ni. si-baya *brother of mother*, Me. baja *brother of father*
- Ni. hɪlua *skin-disease*, Me. belua *leprosy*
- Ni. bute, Me. butet *pointed end of a plant*
- Ni. hɪli-hɪli *uncertain, unsteady*, Me. ele *perhaps*
- Ni. gogo, Me. gugu *lower back*
- Ni. kalamba, Me. kalabba *big boat*
- Ni. dege *approach*, Me. legere *closeness*
- Ni. mi, Me. moi *to come*
- Ni. fili, Me. palau *castrate*
- Ni. savi, Me. sabau *trespass against*
- Ni. tundra *glass-pearl*, Me. tuda *big, long pearl*
- Ni. a-huli, Me. ma-ulau *early in the morning*
- Ni. lave *female*, Me. labai *aunt, elderly woman*

- Ni. balatu *working knife*, Me. balatu
 Ni. huno, Me. enuŋ-an *path*
 Ni. vaha, Me. ban *horn*
 Ni. ajulo, Me. ajolou *egg*
 Ni. momo, Me. meme *loose*
 Ni. lulu *upper end, bed-head*, Me. lulu *to guide, lead*
 Sim., Sich. maila, Me. meira *sea-fish which causes poisoning*
 Sim. la-toru?, Sich. la-tolo?i, Me. turu-turu *alang-alang*
 Sim. maeŋ, Sich. mōēō *hawk*, Ni. moyo *kind of eagle*, Me. mañŋ *eagle*
 Sim. safut-i, Ni. savu, Me. sabu-i *to wipe off*
 Sich. falŋ, Ni. fari, Me. pare *coconut greaves* [left over after oil extraction]
 Sim. əpa, Me. matat kepa, En. e-aro?opa *armpit*
 Sim. atəŋəŋan, Me. tereŋəŋa, En. e-kahaha *scorpion*
 Sim. bai?, Me. bai *just, perhaps*
 Sim. inti?, S.-Me. ta-iti *broken*
 Sim. katuko, Me. katuka *kind of tree*
 Sim. koku? *cohabitante*, S.-Me. koko *husband, wife*
 Sim. -ma?i, Me. -mai *our(excl.)*
 Sim. e-nawan *right side*, En. e-daba *the right one*
 Sim. sibix, Sich. imbi, Ni. simbi *chin*
 Sim. əlis, Sich. əli?, Ni. di *gnat*
 Sim. xexe, kexe, Sich. xexe, Ni. haxi *stalk, stem*
 Sim. bawa, faba, Sich. bawa, Ni. bava *moon, month*
 Sim. batŋ? *chicken enclosure below house*, Sich. batŋ, Ni. batŋ *house*
 Sim. timba-timba *palate*, Sich. timba-timba, Ni. timba *lower chin*
 Sim. tolog, Sich. a-tulŋ, Ni. a-tulŋ *upright*
 Sim., Sich., Ni. tete *back*
 Sim. lahan-laxan, Sich. i-laxa, S.-Ni. salaxa-laxa *guts, heart, stomach*
 Sim., Sich. lixŋ *house*, Ni. liŋu *hut*
 Sim. axisi, ahisi, Sich. axi, Ni. hisi *furious*
 Sim. soŋo fatu, Sich. oŋo, Ni. soŋo *kind of fish*
 Sim. fupub, Sich. a-fufu, Ni. fufu *to reduce to small pieces*
 Sim. atŋ, Sich. fatŋ, Ni. fatŋ *price*
 Sim. daluag, Sich. lalua, Ni. lalua?a *sole, inner part*
 Sim. siŋŋ, Sich. iŋŋ, Ni. siŋŋ *to observe*
 Sim. fusa, Ni. busa *to peel*
 Sim. anan, S.-Ni. hana *why*
 Sim. afasiŋ, S.-Ni. abaso *to burn*
 Sim., Ni. sini-sini *kind of plant*
 Sim., Ni. siŋa *bamboo as a tool*
 Sim. tiŋol, Ni. tibo-?i *to expose*
 Sim. abon, Ni. m-ambu *smith, anvil*
 Sim. iwaŋ, Ni. i?iwa *kind of grass*
 Sim. tiŋa, Ni. tiŋa *basket made of pandanus leaves*
 Sich. mafŋ, Ni. mavi *small wild palm*
 Sich. ufe *look*, S.-Ni. uve *eye*
 Sich. uhu, Ni. susu *to string*
 Sim. kəliŋ, Sich. giŋi *river-mussel*
 Sim. lamon, Sich. lamo *sprouting coconut*
 Sim. kasa = hasa, Sich. xaha *work, feast*
 Sim. tenəŋ, Sich. tini *torch, match*
 Sim. ku = ko, Sich. o-xoxo *kernel, pit*

List 2: Lexical innovations shared by Barrier island languages and Toba-Batak
(and/or other Sumatran languages)

- Sim. tebəl, Ni. simbo, Me. ti(m)bo, En. e-ipo, KBT. simbər *smoke*
 Sim. a-təlu, Sich., Ni. tou, TBt. toru, DaP. tərūh, Ga. tuyuh *under, below*
 Sim. sara, Sich. ala, Ni., Me. sara, TBt., Angk., KBT. sada, Ga. sara *one*
 Sim. alae, ale, Ni. le, Me. alei, TBt., Angk., DaP., KBT. ale-ale *companion, friend*
 Sim. laʔun, Ni. laʔo, Me. lakut, Ga. lakun *brother- or sister-in-law*
 Sim. dəlog, Sich. l+l+l, Me. leleu *hill, forest*, TBt., Angk. dolok, KBT. dələŋ
mountain, Ac. rölōŋ cliff
 Me. ekem, TBt. ehem *to clear one's throat*
 Me. eket, TBt. a-l-hot *sap*
 Me. elak, TBt. holaŋ *space between*
 Me. bukat, TBt. bo-r-gat, bu-r-gat *uproot*
 S.-Me. gude *banana*, TBt. aŋ-gunde-a *banana in the language of the medium*
 Me. pulege, TBt. pulogos *kind of rattan*
 Me. sapo, TBt. sapu *spotted, stained*
 Me. ulup *to blow*, TBt. u-l-tup *to shoot with blowpipe*
 Me. a-kula *flesh*, KBT. kula *body, skin*
 Me. laje, TBt. le, KBT. ləhe *hungry*
 Me. ale, TBt., Angk., DaP. ale *oh*
 Me. belek, TBt. bolon *to fall*
 Me. luŋun, DaP. luŋu *be sad, look for revenge*
 Me. ŋitŋit *mosquito*, TBt. ŋitŋit *moth*
 Me. landrou *limbs*, TBt. lando *length*, KBT. mə-lando *long as of bamboo sections,*
fingers
 Me. om, TBt., Angk. orom, KBT. ərəm *to resist*
 Me. oppat *pull out (from a sheath)*, TBt. uppət, Angk. umpat *to pull out*
 Me. pasi *subterraneous vertical root*, TBt., Angk. pasi *cone*, Ga. pasi *pointed*
end, pin, peg
 Me. suruk-at *pregnant*, DaP. suruŋ *foetus*
 Me. saraina *brother*, KBT. sənina *brother of a man, sister of a woman*, Angk.
 mar-sadaina *have one mother*, Ga. sar-inö *brother, sister*
 Me. sokat, TBt. sogot *next day*
 Me. ale, alei, Lamp. salai *afterbirth*
 Me. kuruk, TBt., Mand. hunduk, Ga. kukuʔ *back, to lie with one's back towards*
 Sim. bəŋiʔ, Sich. bŋgi, Ni. biŋi, KBT., Ga. bəŋkik *bat*
 Sim. muŋkoi, Sich. muŋkui, Ni. mugu, Ga. muŋkus, Ac. muŋkueh *kind of small fish*
 Sim. aŋkix, Ni. ago, TBt., Mand. aŋgo, KBT. aŋgəh *to smell, kiss*
 Sim., Sich., Ni. dalu-dalu, Ga. dədalu *kind of plant*
 Sim. ima mali(x), Sich. imamali, Ni. mali-mali, Mal. mamali *kind of tree*
 Sim. saiŋ, Ni. sai, Min. saiŋ *fang*, Angk. saiŋ *tooth of a horse*
 Sim. abaŋ, Ni. mu-homba, TBt. habaŋ, Lamp. humabaŋ *to fly*
 Sim. oleŋ, Ni. hole-hole, Min. oleŋ *sloping*
 Sim. dan, ran, Ni. a-ra, TBt., Mand. dan *duration, long*
 Sim. tafa, Ni. taba, TBt., Angk. taba, KBT. tabah *to cut, root out*
 Sim. alafae, Ni. alawe, Lamp. kalabay, MML. kələway *female (animal)*
 Sim. pato, Ni. fato, Angk. pato *hatchet*
 Sim. tidao *pray for*, Ni. sindro-a *idol*, Ga. tiro *to ask for*
 Ni. f-al-ea, TBt. p-eak *to lie down*
 Ni. tuo, TBt. tura-tura *to sting*
 Ni. bexu, TBt. begu *spirit*
 Ni. belu, TBt. sidan belu *name of a spirit*
 Ni. fa-biko, TBt. pa-biha *to open*
 Ni. duru-duru, TBt. dolo *kind of shrub*

- Sim. tēpi(x) *piece*, Sich. a-tēpi *a little*, Ni. a-tifi *broken off*, En. e-kopi *piece*
 Sim. tifa, Ni. töva *basket made of pandanus leaves*
 Sim. arin, Ga., Ac. areñ *barb on a spear*
 Sim. riri, Tbt., Angk. didi, Kbt. ridi, Ga. niri *to bathe*
 Sim. baiñ, Kbt. bahñ, Ga. böñ *ginger*
 Sim. baluñ, faluñ, Sich. mbaluñ, Tbt. sibiruñ *heron*
 Sim., Sich. bantae, Min. bantai *flesh, meat*
 Sim. bereñan, Sich. beleñan, Ac. briñan *yard on a sailing boat*
 Sim. kaol, xaol, haol, Tbt., Angk. gaol *banana*
 Sim. gəməto, Ga. gəməto, Ac. gömöto *wasp*
 Sim. hunsa?, xunxa?, Kbt. kunsä, Ga. kunsö, Ac. gunsä *dry measure*
 Sim. lagañ, Sich. ilaxañ, Angk., Ac. lagañ *kind of tree*
 Sim. abui, Min. abuih *to cook in water*
 Sim. boruñ, Sich. oluñ, Ac. burōñ *demon, spirit of a dead person*
 Sim. ana?, Sich. g-ana?, Ga. anas *prepared betel*
 Sim., Angk. nali, Kbt., Ga. nalih, Ac. naleh *rice measure*
 Sim. saeam bano, Ga. sayam, Angk. sayom, Tbt. saem *to bring back to harmony*
 Sim. dabıs, Ga. döböş, Ac. daböeh *ware, article*

4.5 Semantic innovations shared among the six languages under investigation

Further evidence for our subgrouping hypothesis is found in the following lists of semantic innovations which appear to be exclusively shared.

List 1: Semantic innovations shared by Barrier island languages

- Sich. falı, Ni. balı, Me. bale *to borrow* (< *bales *to repay*)
 Sim. bano, En. e-pado *placenta* (< *banua *land, settlement*)

List 2: Semantic innovations shared by Barrier island languages and Toba-Batak (or other Sumatran languages)

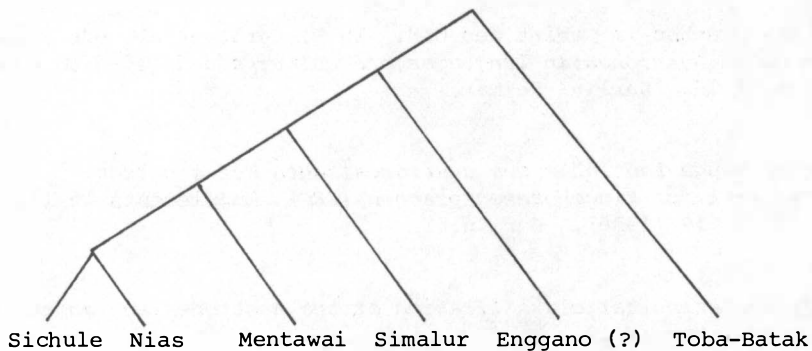
- Me. ulou, Tbt. ulok *snake* (< *qulej *worm, maggot*)
 Me. tuktuk, Tbt., Kbt. t-ar-utuñ *Durian* (< *tu(η)tuñ *spinous animal*)
 Me. paola, poula, Tbt., Kbt., DaP., Ga. pola *sugarpalm* (P-Minahassa *pola *sugarcane*)

4.6 Phonological irregularity and semantic innovation shared by Barrier island languages and Toba-Batak (and/or other Sumatran languages)

- Sim. ləkao *dry season*, Sich. lixiñ *heat which follows rain*, Ni. lixiñ *clear (of weather)*, Me. ma-legeu *warm, dry (of weather)*, Tbt., Angk. logo, Kbt. ləgo *dry (of weather)* (< *qalejaw *day*)

4.7 Internal relationships of the Barrier island-Batak group

Considering the number of phonological innovations exclusively shared among members of the Barrier island-Batak group one might suggest the following tentative internal subgrouping:



NOTES

1. This is a slightly revised version of a paper presented at the Fourth International Conference on Austronesian Linguistics at Suva in 1984. I thankfully acknowledge the helpful comments of Robert A. Blust, David Zorc and S. Adelaar.
2. Abbreviations used in the body of the paper: Ac. = Achinese, Angk. = Angkola-Batak, DaP. = Dairi Pakpak, Dou. = Douru, En. = Enggano, Ga. = Gayo, KBt. = Karo-Batak, Lamp. = Lampung, Mand. = Mandailing, Min. = Minangkabau, Me. = Mentawai, MML. = Middle Malay, Ni. = Nias, Sich. = Sichule, Sim. = Simalur, TBt. = Toba-Batak.
3. The linguistic evidence for metal, rice and weaving is discussed in Blust 1976.

REFERENCES

ADRIANI, N.

- 1928 Spraaakunstige schets van de taal der Mentawai-Eilanden. *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch-Indië* 84:1-117.

BLUST, Robert A.

- 1976 Austronesian culture history: some linguistic inferences and their relations to the archaeological record. *World Archaeology* 8/1:19-43.
- 1978 Eastern Malayo-Polynesian: a subgrouping argument. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, Fascicle 1, *Western Austronesian*, 181-234. *PL*, C-61.
- 1981 Variation in retention rate among Austronesian languages. Paper presented at the Third International Conference on Austronesian Linguistics, Denpasar.

CAPELL, Arthur

- 1982 Bezirkssprachen im Gebiet des UAN. In R. Carle et al, eds *Gava^c: Studies in Austronesian languages and cultures dedicated to Hans Kähler*, 1-15. Berlin: Reimer.

DEMPWOLFF, Otto

- 1934-38 *Vergleichende Lautlehre des austronesischen Wortschatzes. Zeitschrift für Eingeborenensprachen (ZES)*, Supplements 15 (1934), 17 (1947), 19 (1938). Berlin.

DYEN, Isidore

- 1965 *A lexicostatistical classification of the Austronesian languages. IJAL* 31/1, Memoir 19.

ESSER, S.J.

- 1938 Talen. In *Atlas van tropisch Nederland*. Amsterdam.

JONKER, J.C.G.

- 1918 Taal. In *Encyclopaedie van Nederlandsch Oost-Indie*. The Hague: Leiden.

KÄHLER, Hans

- 1937 *Untersuchungen über die Laut-, Wort- und Satzlehre des Nias. ZES* 27. Hamburg, Berlin.
- 1940 Grammatischer Abriß des Enggano. *ZES* 30/2-4:81-117, 182-210, 296-320.
- 1959 *Vergleichendes Wörterverzeichnis der Sichule-Sprache auf der Insel Simalur an der Westküste von Sumatra*. Veröffentlichungen des Seminars für Indonesische und Südseesprachen der Universität Hamburg, 1. Berlin.
- 1961 *Simalur-Deutsches Wörterbuch mit Deutsch-Simaluresischem Wörterverzeichnis*. Veröffentlichungen des Seminars für Indonesische und Südseesprachen der Universität Hamburg, 3. Berlin.
- 1975 *Texte von der Insel Enggano (Berichte über eine untergehende Kultur)*. Veröffentlichungen des Seminars für Indonesische und Südseesprachen der Universität Hamburg, 9. Berlin.
- n.d. Grammatik der Simalursprache. MS.

LAFEBER, Abraham

- 1922 *Vergelijkende klankleer van het Niasisch*. The Hague.

MAHDI, Waruno

- n.d. Morphophonologische Besonderheiten und historische Phonologie von Malagasy. MS.

MARSCHALL, Wolfgang

- 1966 Sind die Kulturen von Mentawai altindonesisch? *Paideuma* 12 = *Mitteilungen zur Kulturkunde* 1:128-134.

MORRIS, Max

- 1900 *Die Mentawai-Sprache*. Berlin.

SALZNER, Richard

1960 *Sprachenatlas des Indopazifischen Raumes*. Wiesbaden: Harrassowitz.

SCHEFOLD, Reimar

1972 Divination in Mentawai. *Tropical Man* 3:10-87.

1976 Religious involution: internal change and its consequence in the taboo-system of the Mentawaians. *Tropical Man* 5:46-81.

1979/80 *Speelgoed voor de zielen. Kunst en cultuur van de Mentawai-Eilanden*. Delft/Zürich.

SUNDERMANN, H.

1905 *Niassisch-Deutsches Wörterbuch*. Moers.

TUUK, H.N. van der

1897- *Kawi-Balinesesch-Nederlandsch Woordenboek*. 4 vols. Batavia.
1912

WARNECK, J.

1905 *Tobabataksch-Deutsches Wörterbuch*. Batavia.

WILLMS, Alfred

1955 *Lautliche und syntaktische Untersuchungen über die Mentawai-Sprache. Afrika und Übersee* 40. Hamburg.

WURM, S.A. and Shirô HATTORI, eds

1983 *Language atlas of the Pacific area, part 2: Japan area, Philippines, Taiwan, mainland and insular south-east Asia*, in particular map 38, with W.A. Foley's notes. Canberra: The Australian Academy of the Humanities in collaboration with the Japan Academy; PL, C-67.

ZAINUDDIN HR LENGANG, BE KIM HOA NIO, MOHD. ANSYAR ZAINIL, and SYOFYAN ADAM

1978 *Bahasa Mentawai*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.

SOME PROBLEMS OF DIACHRONIC TYPOLOGY OF THE MALAYO-JAVANIC LANGUAGES

A.K. Ogloblin

The aim of the present paper is to summarise data on the historical morpho-phonology of some languages of the Java-Sumatra-Malacca Peninsula area (see Wurm and Hattori 1983, especially maps 38 and 39) and to analyse these data from the point of view of areal typology. The languages under observation are mainly Malay (Mal), Sundanese (Snd), Javanese (Jav), and Madurese (Mad), and in part also Acehnese (Ac) and Minangkabau (Mkb). We shall term these the Malayo-Javanic languages without ascribing to the term any strictly genetic meaning.¹

In analogy to synchronic typology we can assume the aim of diachronic typology to be the classification of languages according to types of structural changes and, further, the establishment of logical relationships between those changes. If such a problem is limited areally, we are often able to present different typological classes of languages as successive stages of a single process or of a certain continuing historical trend (see for instance Sharadzenidze 1982:44). Thus, Yakhontov (1971) writes of the "Sinitic" languages of China and South-East Asia: "The development of all isolating syllabic languages takes one and the same direction ... we ... can trace a common trend, a common direction, comparing analogous changes taking place in different, often unrelated languages". According to these common trends the author subdivides the Sinitic languages into archaic, middle (intermediate) and late, depending upon root morpheme changes, presence of tones, composition as a kind of word-formation, etc., with a class not necessarily comprising languages co-existing in time, for example Khmer and Old Chinese as members of the archaic class. Another example of this is Hutterer's classification of the Germanic languages (Hutterer 1970). He divides all modern languages of this group, from Icelandic to Afrikaans, into five classes placed along a "modern-archaic" axis, and here again, one typological class may include languages of different epochs, as for instance Icelandic and Old English.

It has been pointed out, concerning the Austronesian languages, that the greatest number of archaisms is to be found in the north of the area over which they are spread, including Taiwan and the Philippines (Sirk 1978). Dahl and other authors present the PAN verb system as very close to that of the Philippine languages (Dahl 1976). The Malayo-Javanic languages thus become included in a wide zone of structural innovations. It is also interesting to mention in this context that, as had already been observed in the 19th century, Old Javanese shows a greater likeness, materially and structurally, to the Philippine languages than does modern Javanese. For example, the infixes (infix-prefixes) -in-, -um-, characteristic of the Philippine languages, are productive in Old Javanese (the

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 111-122. *Pacific Linguistics*, C-94, 1986.

© A.K. Ogloblin

111

second of them occurs in the sparse Old Malay inscriptions, too) but have either disappeared or lost their productivity in the modern languages. A characteristic Philippine trait - insertion of the infix into the prefix: paka- - p-in-aka- is known in Old Javanese, but has disappeared in New Javanese (together with the prefix paka-).

One of the reasons for innovation in the Austronesian family, just as in other languages of the world, is contact with languages of different structure. One can name as examples Cham and the other AN languages of Indochina existing surrounded by isolating languages unrelated to them genetically, the languages of East Indonesia with a supposed non-AN ("Papuan") substratum, Malagasy with a Bantu substratum. Language contacts, however, present a great number of various linguistic situations that can have very different consequences in evolution. The contact may possibly sometimes accelerate a trend in the inner structure of the system. This may be the explanation of parallel development in languages that are remote territorially but close typologically. The importance of the problem of distinguishing between internal and external factors of development of a language system and between both these factors and the universal factor of the functional improvement of a language is obvious, and does not need any comment.

As for classification markers (distinctive features), they can be more or less close to the features of a global or universal typological classification, depending partly on the language level taken for observation. Ye.M. Wolf, summing up her comparative study of the Ibero-Romanic languages, wrote: "Typological features worked out in general typology cannot be used for languages so close structurally" (Wolf 1970:246). Indeed, from the point of view of "macrotypology" such languages fall into one and the same class, inside which division is carried out on a different basis from the delineation of the class itself. But in morpho-phonology areal classification features are, on the whole, less diversified. Some of the features given below concern the Malayo-Javanic group as a whole and are of general significance.

CHANGES IN THE MORPHEME SYLLABIC STRUCTURE

In morpheme syllabic structure the Malayo-Javanic languages are classed together as languages with a predominance of disyllabic morphemes and are subdivided by the proportion of monosyllabic ones. Sundanese, according to Fokker (1953:22), has no less than 80% disyllabic morphemes. In Madurese the situation seems to be much the same. Javanese has 85% (Uhlenbeck 1949:23) and Indonesian 87% (Shchuko 1973:174) disyllables. Disyllabism is, judging from the Dempwolffian PAN root stock, where it amounts to 96% (Gonda 1950:323), a historically stable characteristic of AN languages.

There are few monosyllabic morphemes (less than 5%) in Javanese, Indonesian, Sundanese and Madurese (for the first two, see Uhlenbeck 1948 and Shchuko 1973). Zubkova (1977:213) notes significant alternation of the non-typical monosyllables with the typical disyllables in Indonesian: the latter are formed by means of a prothetic vowel /ə/: /gʊŋ/ - /əgʊŋ/ *gong*. This prothesis is characteristic, too, of other languages of Java: Mad əjjham /əcc^cʌm/ - Mal /jam/,² /əppɔt/ - Dutch *pot* *pot*, *vase*, Snd /ərcɔs/ - Dutch roots *rose*, Jav (variants) /əlɛr/ - /lɛr/ *north*. Another way of disyllabisation may be "splitting" of a vowel, its distribution into two syllables with a consonant element inserted: Snd /gɔʔɔŋ/ - Jav /gɔŋ/ *gong*, Snd /saʔat/ - Jav /a-sat/ *dry*, Mad /rɔʔɔm/ - OJav rūm *aromatic*, Mkb Ruhum -

Ar *rūm Byzantium; Turkey*, Ac *teu'ōt*, Mad /təʔət/, cf. Mal *lu-tut knee* and *ber-tel-ut kneel* (see Fokker 1953:40-41; Toorn 1891:ix; Cowan 1948:432).

In Old Javanese there occur not infrequently monosyllables that are reconstructed as PAN disyllables and have undergone vowel contraction OJav *doh* < PAN *d'auh *far*, OJav *woh < PAN *buah *fruit*, OJav *bot* < PAN *bəyat *weight*. Later these monosyllables turned again into disyllables (being either replaced or retained as variants), so that a portion of Javanese vocabulary has undergone a cyclic process from disyllable to monosyllable and back to disyllable. The reverse disyllabisation is treated in considerable detail by J. De Casparis. The main role here belonged to prothesis of the vowels /ə/, /i/, /u/, affixation, reduplication and morpheme merger with subsequent lexicalisation of the derived disyllable: *woh* - /uwəh/ *fruit*, *tūt* - /ətət/ *follow*, *srah* - /pasrah/ *give away*, *gya haste* - /gege/ *hasten, hurry*, *tan wruh* - /tambəh/ *ignorant of*. "Where influence of phonetic laws threatened the disyllable structure" - the author says - "the speakers have found ways of restoring it" (Casparis 1947:76).

Monosyllables derived from former disyllables are also numerous in Acehnese and in the Malay Peninsular dialects. In Acehnese, as Cowan points out, a considerable part of the word stock comprises monosyllables corresponding to disyllables in cognate languages: Ac *thon* - TM *tahun year*, Ac *tron* - TM *turun descend*, Ac *ba* - TM *bawa carry* (Cowan 1974:200-203). In the Peninsular Malay dialects monosyllabisation has been pointed out by different authors (see data in Ismail 1973).

In Acehnese and the Malay dialects monosyllabisation is probably caused by contact. The considerable material and structural similarity of Acehnese to the Mon-Khmer languages, treated in detail by Cowan, led to supposing a Mon-Khmer substratum, though the author tends rather to stress the contact of Acehnese with Malay.³ In western Malaysia monosyllabisation is most prominent in the areas of contact with Thai and, in the past, with the Mon-Khmer languages (see also Ogloblin 1983b). It is difficult to find an explanation for the monosyllabisation found in the Old Javanese vocabulary; it is possible that in ancient times here too some contact took place, traces of which are lost in history.

SYNCHRONIC VARIATION OF MORPHEME SYLLABIC STRUCTURE

Reduction of the syllabic structure of meaningful elements in fluent speech probably exists in all languages, but rules within this trend are not universal for languages of different types. In the languages under discussion three types of such general rules are known.

1. Shortening to the final syllable: Snd *lamun* - *mun*, Mad /lamən/ - /mən/ *if*, In /satu/ - /tu/ *one*, /ibu/ - /bu/ *mother*, Jav /maju/ - /ju/ *forward*, /pəḍə/ - /ḍə/ *the same, no matter* (Uhlenbeck 1949:65), Mls /apə/ - /pə/ *what?*
2. Contraction of contiguous vowels or vowels divided by a fricative sonant (sonoric continuant): Jav /duwe/ - /de/ *have*, /mau/ - /mu/ *lately* (Uhlenbeck 1949:59), Snd /naha/ - /na/ (Fokker 1953:22); here belongs, also, the diphthongisation of a disyllabic vowel group: In /mau/ - /məy/ *wish, intend*, /məmulai/ - /məmulaj/ *begin* (Zubkova 1977:215).
3. Omission of a vowel with retention of the preceding consonant: In /bəri/ - /bri/ *give*, Ulu Muar /təpi/ - /təpi/ *edge*, təmpayan/ - /təmpayan/ *jug* (Hendon 1966: 109-110).

The first of these rules sets apart Peninsular Malay (dialects and Malaysian) where shortening is widespread. The second rule is practically unexplored. The third rule has a very clear diachronic projection, making it possible to divide the languages under study according to the stage of development of this alternation. Four such stages can be set out, including the initial and the final ones, where alternations are nearly or completely non-existent.

STAGE 1: Alternation with a material zero practically does not occur: OJav, Snd and Mad. In the OJav vocabulary there are few variants: *geremus* - *gremus* *scratch*. In the texts the vowel /ə/ was often omitted, but these omissions, judging by their correspondence to the rules of metrical versification, did not reflect pronunciation (Kern 1871:2). In literary Snd and Mad alternation occurs under the influence of borrowing from Jav, which is especially evident in the Western Mad dialect, closer to Jav than the literary language: Mad /carəta/ - /cɛta/ *story*, Snd /karana/ - /krana/ *cause*. The Snd explanatory dictionary (Satjadibrata 1950) often gives reference to the full-vocalised variant.

STAGE 2: In antepenultime the non-neutral vowels /a/, /i/, /u/ alternate with the neutral shorter /ə/, and the neutral vowel in antepenultime and penultime alternates with zero: Jav and Mal, except for some Peninsular dialects: Jav *nagara* /nagərɔ/ - /nəgərɔ/ (*royal*) *city*, /sinuhon/ - /sənuhon/ *ruler, potentate*, /gumagos/ - /gəmagos/ *smart* (see, with different transcription, Uhlenbeck 1949:14), /mə laku/ - /mlaku/ *walk*, /gaməlan/ - /gamlan/ *gamelan* (Horne 1961:xxix), In /kalimat/ - /kəlimat/ *sentence, phrase* (Alieva et al 1972:29), /pərau/ - /prau/ *boat*, Mls /dərama/ - /drama/ (the second variant is considered more correct at present). Because of the complex character of this distinctive feature it permits scope for exploring in greater detail than is considered here.

STAGE 3: Other non-neutral vowels alternate with the zero. This can be observed in the Ulu Muar Peninsular dialect of Malay: /sakɪt/ *ill* - /skɪt atɪ/ *offended*, /situ/ - /stu-a/ *there* (Hendon 1966:109-110) (the latter variants with special intonation).

STAGE 4: Alternation is absent: the former zero is now simply a boundary between abutting consonants. To this stage belongs Acehnese, and among the Peninsular Malay dialects, the speech of the inner regions of Kedah: *thong* - TM *tahun year* (Ismail 1973:76) (examples are sparse).

ASYLLABIC MORPHEMES

Morphemes containing no syllabic elements, that is practically vowels, are extremely few in the Malayo-Javanic languages. They all are function (auxiliary) morphemes, and most of them have two allomorphs - a syllabic and an asyllabic one, sometimes one of them being stylistically marked: either the syllabic one is bookish, as Jav aN-, or the asyllabic one is colloquial, as In N-, cf. *mukul* (= N-pukul) instead of the literary *memukul* (= /məN-pukul) *beat*. From the point of view of typology asyllabism of relational morphemes seems to be more important than that of derivational ones: the latter often are unproductive, and the corresponding derivatives are more or less simplified. Among the relational morphemes, the Old Javanese article /ŋ/ should be noted and the actor person markers found in some texts (Teselkin 1963:45 and 59). These morphemes have disappeared in New Javanese. On the other hand, unlike Old Javanese aN-, in the languages of Java and partly in Malay dialects a verbal prefix N- is used,

realised either segmentally - as one of the nasal sonants - or appearing as a distinctive feature of nasality in a consonant, belonging in a way by fusion both to the prefix and the root or base: Mad /ŋ-ɔwan/ *herd*(v.), Jav /n-donɛŋ/ *tell a fairy-tale*, compare Mad /maca/ - /N-baca/ *read*, Jav /nrimɔ/ - /N-trimɔ/ *submit*. In TM and literary In and Mls there are no asyllabic relational morphemes; for the Peninsular dialects data are few: we can point out a pronoun-anaphoric element /n/ in Ulu Muar and a proclitic morpheme /s/ *one* in Kedah in the position before a root-vowel: /saRi/ *one day* from /aRi/ *day*. In Snd and Mad asyllabic morphemes besides the prefix N- are unknown, at least not productive ones, except the /-n/ allomorph of the suffix /an/ in Mad: /k^cibʌ/ *carry* - /bʌn-k^cibʌn/ *burden*.

All in all the feature of morpheme asyllabism is difficult to use for diachronic classification of the languages under observation. However, it has great importance for general typology. As Kasevich (1983) shows, absence of asyllabic morphemes is characteristic of syllabic languages. Thus, according to this feature, the Malayo-Javanic languages prove to be close to the syllabic type.

SEGMENTAL STRUCTURE OF MORPHEME AND SYLLABLE

Under this heading fall the features of a) initial consonant clusters in morphemes and syllables, b) consonant membership in the final position, c) intervocalic consonant clusters in a morpheme. The first feature is connected with the variability of syllabic structure discussed above. Omission of the vowel leads, as can be seen in the examples given above, to formation of initial consonant clusters. In Snd and Mad the tendency to avoid such initial groups is stronger than the tendency towards disyllabism, compare mutation of loanwords: Mad /kalebun/ - Jav /kliwɔn/ *headman*, Snd Parasman - Dutch Fransman *Frenchman*. In Javanese the picture is reversed. Already in Old Javanese initial groups of consonants in loanwords as a rule did not assimilate, though in the original root stock such groups were utterly or nearly non-existent. Later the OJav trisyllables, losing the first vowel before the liquids /r/, /l/, formed NJav disyllables: kalambi - /klambi/ *coat, shirt*, kulimis - /klimes/ *smooth, slippery*, kirincing - /kriŋcɛŋ/ *ringing*. In Malay at an early stage, as evidenced by loanwords of the type of seloka *poetic maxim* - Sanskrit çloka *an epic metre*, initial consonant clusters were avoided, whereas in Indonesian and the Jakartan dialect they are normal (Grijns 1981). In Malaysian such groups are, as we may observe, also quite frequent, and codified by the latest spelling rules as well. They alternate regularly with segments containing a vowel, usually /ə/. The same is to be found in the Peninsular dialects, with the main variant being one either with or without the vowel. In Acehnese consonant clusters at the beginning of a morpheme are quite numerous and varied.

The feature of limitation of consonantism at the end of the morpheme gives three typological classes.

1. Old Jav and Snd, where voiced consonants are possible not only in the *anlaut*, but also in the *auslaut* (compare Tagalog).
2. Mal, Jav and Mad. In the former voiced consonants in final position are found only rarely in loanwords before the initial vowel of a suffix: /jawab-an/ *answer* (variant /jawapan/). In Madurese final voiced sounds also do not occur; final aspirates appear regularly before the suffixed vowel: /tɔtɔp^c-i/ *close*(v.) In Javanese pronouncing aspirates in this position is obsolete (Uhlenbeck 1949:42-43). Still final consonant position puts Mad close to OJav and Snd, as Mad voiced

consonants can occur in the *inlaut*: /kəḍ/ḍiʔ/ *worry* (in rare cases also possible in In with juxtaposition of different stops: /ab/dul/ (compare Stokhof 1980:43).

3. Acehnese and Malay Peninsular dialects, where the majority of consonants do not occur finally. Sometimes a sound absent in *auslaut* may be restored before a suffix: Ulu Muar *tapih filter*(v.) - *tapis-an filter*(n.) (Hendon 1966:24). Such cases are caused mostly by contact with the literary language (compare further comments on Mkb).

As for the *inlaut* of a morpheme, we can look for a complex feature of "the degree of variability in possible intervocalic consonant groups". But practical use of this feature is complicated by the fact that standard languages have a great number of rare consonant combinations in non-assimilated or partly assimilated loanwords; e.g. in Mad the cluster /str/ is found in two or three words only, as *éstré woman* (polite style) < Sanskrit *strī*. Excluding loanwords, two directions of innovations can be outlined: towards the increase and the decrease of diversity in intervocalic clusters. The first tendency is observable in Mls (standard and especially dialects). The Kelantan dialect has lost the homorganic groups "nasal sonant - voiceless stop": Kel *laca* - TM *lancar fluent, rapid* (Asmah 1977). In most Peninsular dialects and in the prevailing pronunciation of standard Mls, omission of the TM -r at the end of a syllable simplifies the corresponding intervocalic combinations: /kəjə/ - TM *kerja work*. The second direction of innovation - towards a greater diversity of intervocalic groups - is characteristic of Mad. Intervocalic gemination is - at least for a considerable part of the vocabulary - an innovation; compare loanwords: Mad *pojja* < Skr *pujā reverence*, Mad *nabbi* < Ar *nabiʔ prophet*, Mad *rassa* < Skr *rasa emotion*, Mad *raddhin* < Jav *radèn* (a title of nobility). Besides that, Mad has developed combinations "nasal stop - aspirate": *təngghi*, cf. Mal /*tingi*/ *high*.

Jav Old and New, Snd and In are intermediate with respect to this feature, though a more detailed study reveals distinctions between them. Ac and Mkb are close to the Peninsular type, as intervocalic clusters are avoided here: Ac *meuseukin*, Mkb *misikin* or *mikin* - Mal *miskin poor* (from Ar).

SUPRASEGMENTAL PHONOLOGY

The data for classification within the Malayo-Javanic group according to this feature are rather insufficient; however, the situation among other related languages is interesting. Unlike such languages as Tagalog, Buginese or Toba Batak, it is typical of a morpheme or word in Malayo-Javanic languages to have no phonologically relevant prosodic or other phenomena, including word stress. A possible exception may be the nasal assimilation of vowels in Snd discovered by Robins (1957), that seems to distinguish different morpheme structures (though minimal pairs are not found). In other languages suprasegmental elements serve to segment the text (Trubetskoy's delimitative function). They include vowel harmony with respect to tongue height (closure) in In, possibly also in Jav and Snd. This harmony is limited to the root morpheme (on In see Emeis 1955 and Zubkova 1970). In contrast to this in Mad there is a progressive assimilation of vowels with respect to closure, involving not only the root morpheme but also the suffix (Stevens 1968). Absence of stress with a phonological function must be an innovation of the Malayo-Javanic languages as a whole; appearance of such stress as an innovation in the Philippine and other cognate languages is less probable.

As for delimitative stress, it does not seem to be quite the same in the different languages of the group. In particular, diphthongisation of vowels in the final syllable in Mkb, Peninsular Mal and Ac may, according to some authors, Cowan among them, be linked with stress movement (delimitative, not phonological) to this syllable. For all these languages it coincides with elimination of the first syllable in part of the vocabulary (see above; on Mkb see Kähler 1965).

TENDENCY OF COINCIDENCE OF MORPHEMIC AND SYLLABIC BOUNDARIES

Total coincidence of the two sorts of boundaries is a law of syllabic languages such as, for instance, Chinese or Burmese. Historical changes in correspondence between syllable and morpheme boundaries mean either a drift towards syllabic type, if boundary coincidence increases, or away from it, if this coincidence becomes more rare. In the phonological typology laid out by Kasevich (1983) some Indonesian languages are classified as asyllabic ones with syllabic features. Their distinction from the syllabic languages is resyllabisation, i.e. the shift of syllabic boundaries against morphemic ones: the consonant of one morpheme falls into the same syllable as the vowel of another: In /minu/ *drink* (v.) - /minu/m-an/ *drink*(n.). In syllabic languages resyllabisation can never occur.

In the Malayo-Javanic languages resyllabisation may be prefixal, infixal and suffixal, according to the morphemes that the preceding consonant and the following vowel belong to: prefix and root, as in In /ikut/ *follow* - /pə/ŋ-i/kut/ *follower*; root, infix and root again, as OJav tulis *write* - t-i/n-u/lis *written*; root and suffix (In example above).

The sphere of resyllabisation is accordingly narrowed, other things being equal, if a language loses the productivity or decreases the number of 1) infixes with initial vowel and/or final consonant, 2) prefixes with final consonants, 3) suffixes with initial vowels, 4) roots with final consonants, 5) roots with initial vowels. Here syllabic ("no shorter than a syllable") morphemes are meant, and if allomorphs exist, the ones that count are those that take part in affixation.

All the listed phenomena occur in the Malayo-Javanic languages.

1. Infixes of the VC type are productive in OJav and Snd: Snd /bɔdɔ/ *stupid one* - /b-ar-ɔdɔ/ *stupid ones* (Pavlenko 1965:39). In Njav the infix -um- is unproductive, and -in- belongs to bookish style. Infixes are unproductive in both Mal and Mad, but judging by isolated examples from ancient epigraphy and some simplified morphemes, the infix -um- existed in Old Mal too.
2. The OJav prefix aN- is reflected in Njav N- (except in bookish style, cf. above). In TM, In and Mls the final nasal sound of the prefix /məN-/ passes into the next syllable: /asah/ *sharpen*, *grind* - /mə/ŋ-a/sah/ (active form). In Ulu Muar and Mkb in the position before the root vowel the nasal element is generally dropped: Ulu Muar /məsasah/, Mkb /maaja/ *teach* - TM /mə/ŋ-a/jar/.
3. The suffixes with initial vowel -i and -an in the Mal dialects are less productive than in literary Mal. In Ac all suffixes have disappeared. In Jav the suffix /ake/ is often pronounced /ke/, and sometimes spelt ke: nerusake - neruske *continue*, nguntungake - nguntungke *be favourable*.
4. Final TM -r and -l are lost in Mkb and in some Peninsular dialects; in some of the latter, other consonants as well. Thus, TM ambi/l-an *something taken*

corresponds to Mkb /ambi_ɛ/an/ *tree with leaves torn off*, TM mengai/r-i *irrigate* - Mkb maai_ɛ/i/ (yet the final -r or -l of the root is often restored before a suffix; the same occurs in Peninsular Mal). In Mad part of the former voiceless root finals are replaced by a glottal stop that does not take part in resyllabification: Mad /tɔɔʔ/ - Mal /turu/t-i/ *follow*. We seem to have the same syllabification in Jav, where initial glottalised plosive is not generally found: /tindaʔ/an/ *step*. The same, evidently, in Mkb: /laujʔ/an/ *ocean* - TM /lau/t-an/.

5. Replacing the root with initial vowel by one with initial consonant occurred in the history of Jav with the loss of initial vowel in trisyllables: /yuyu/ < ayuyu *crab*, /tɔɔ/ < Skr *uttama highest, supreme* (sometimes with variants retained). In Mad a number of vowel-initial roots take a formant k-: oca' /ɔcaʔ/ *speech* - /ɛkɔcaʔ/ *to be spoken*.

The facts listed above are evidence that at an early stage of Jav and Mal resyllabification, leading to discrepancy of morpheme and syllable boundaries, was more common than at present. In Ac, Mal Peninsular and Mkb the factors for resyllabification are comparatively weakened, and Snd, as in other respects, is more conservative, being closer to OJav. In Mad, gemination before a vowel-initial suffix is a sort of step back towards the OJav type, where such gemination is presumably a relic of a formerly regular phenomenon,⁴ e.g. k-in-on/n-akan *is ordered* (root *kon to order*), p-in-agəh/h-akan *is affirmed* from the root *gəh firm* (examples from a 10th century inscription (see Sarkar 1972:61)). The second consonant of the geminate belongs to the root, but falls into the same syllable as the suffix vowel. In Mad this is regular: /tɔɔn/ *descend* - /tɔɔn/n-ak^{ci}/ *lower, cause to descend*. The general tendency of the languages discussed with regard to this feature is towards the syllabic type.

CONCLUSION

Tendencies towards shortening (reducing) of the morpheme syllabic structure, towards coincidence of syllabic and morphemic boundaries, towards contrastiveness of consonantism in different positions in morpheme and syllable, a smaller number of asyllabic morphemes (and their alternating with syllabic ones), absence of phonological stress - all this brings the Malayo-Javanic languages closer to the syllabic and isolating languages of mainland South-East Asia. Nor is this contradicted by the formation of new initial consonant clusters in morpheme and syllable - such clusters are noted in the archaic "Sinitic" languages, but disappear at a later stage.

Measured against the common tendencies of the Malayo-Javanic group, more archaic and more advanced languages in that group may be pointed out. The archaic ones, along with the chronologically early OJav (and Old Mal) comprise Snd and, to a slightly lesser extent, Mad. The morphophonological innovations of Mad - increased complexity of intervocalic clusters, including gemination, have a peculiar "archaising" quality. To the more progressive ones belong Ac and the Peninsular dialects of Mal. Literary Mal (In and Mls) and Jav are intermediate as to the general tendencies of the group. Mkb is evidently very close to the advanced group (if not a member of it).

If the general tendencies of the "progressive (advanced)" languages are strongly marked, that may be the effect of contact with a language of isolating and syllabic (or close to syllabic) type. For a language like Jav that has no such contacts areal "long range" effect may be suggested, or we can take the view that contacts simply accelerate the innate evolutionary inertia of the group. These conclusions and assumptions may find confirmation through investigation of data from other languages of the same group and other language levels.

NOTES

1. The Malayo-Javanic group of languages as a genetic unit comprising the first four of the languages listed above is established in Nothofer 1975. The genetic limitations of the group based mostly on the strength of lexicostatistics are criticised in Blust 1981. I am inclined to give the term "Malayo-Javanic" rather a broad interpretation. Of all the varieties of Malay, with a total number of about 15 million speakers, we shall consider the traditional Malay (TM) of medieval literature, Indonesian (In) and Malaysian (Mls) as two modern forms of literary Malay, and the Peninsular dialects. Minangkabau (5.6 million, West Sumatra), judging by its Arabic script that could not have originated before the 16th century, must have become differentiated from Malay at a comparatively recent date. The number of speakers of the languages of Java and Madura is: Jav 67 million, Snd 19 million, Mad 8.7 million, of Ac in North Sumatra 2.2 million as estimated for the end of the 1970s in Bruk 1981 (other figures above are from the same source).

A survey of contacts in the Javanese languages area is contained in Ogloblin 1983a.

2. In the Javanese explanatory dictionary (Poerwadarminta 1931) are found frequent references from a disyllable with /ə/ to the monosyllable, the latter being evidently a more literary variant, sometimes also more close to the source language word. In Madurese the prothesis /ə/ is followed by gemination of the initial consonant of the monosyllable.

Here and further c and j mean palatal stops, y means palatal sonant (fricative).

3. For a contrary view see C. Snouck-Hurgronje: "An ordinary Acehnese not living by the sea-coast, and even most port-residents know hardly a word of Malay" (Snouck 1892:14).
4. See, e.g. Ras 1968. Nothofer (1975) projects gemination into "Proto-Malayo-Javanic". For some ideas on the origin of Mad gemination see Dahl 1981.

ADDENDUM

Of course, the problem of initial consonantal clusters in OJav deserves special investigation. Still the new dictionary that I shall be completing after having finished this paper, gives many instances of CC... groups being referred to CVCV... entries. See Zoetmulder 1982.

REFERENCES

- ALIEVA, N.F., V.D. ARAKIN, A.K. OGLOBLIN and Yu.H. SIRK
1972 *Grammatika indonezijskogo jazyka*. Moscow: Nauka.

ASMAH HAJI OMAR

- 1977 *The phonological diversity of the Malay dialects.* Kuala Lumpur: Dewan Bahasa dan Pustaka.

BLUST, R.

- 1981 The reconstruction of Proto-Malayo-Javanic: an appreciation. *BijdrTLV* 137/4:456-469.

BRUK, S.I.

- 1981 *Naselenie mira: etnodemograficheskiy spravochnik.* Moscow: Nauka.

CASPARIS, J. De

- 1947 L'importance de la disyllabie en javanais. *India Antiqua*, 63-76. Leiden.

COWAN, H.K.J.

- 1948 Aanteekeningen betreffende de verhouding van het Atjehsch tot de Mon-Khmer-talen. *BijdrTLV* 104:429-514.
- 1974 Evidence of long vowels in early Achehnese. *Oceanic Linguistics* 13/1-2:187-212.

DAHL, O.Chr.

- 1976 *Proto-Austronesian.* Lund: Studentlitteratur.
- 1981 *Early phonetic and phonemic changes in Austronesian.* Oslo: Universitetsvorlaget.

EMEIS, M.G.

- 1955 Vocaal-harmonie in de Maleise kern van de Indonesische vocabulaire. *BijdrTLV* 111:191-201.

FOKKER, A.A.

- 1953 Tatabunji Sunda. *Bahasa dan Budaja* 1/6:15-28.

GONDA, J.

- 1950 Indonesian linguistics and general linguistics. *Lingua* 2/3:308-339.

GRIJNS, C.D.

- 1981 Jakartan speech and Takdir Alisjahbana's plea for the simple Indonesian word-form. *Papers on Indonesian languages and literatures*, 1-34. London, Paris.

HENDON, R.S.

- 1966 *The phonology and morphology of Ulu Muar Malay (Kuala Pilah district, Negeri Sembilan, Malaya).* New Haven.

HORNE, E.C.

- 1961 *Beginning Javanese.* New Haven, London.

HUTTERER, C.J.

- 1970 Über die Möglichkeit einer historischen Typologie der germanischen Sprachen. In *Theoretical problems of typology and the Northern Eurasian languages*, 171-176. Amsterdam.

ISMAIL HUSSEIN

1973 Malay dialects in the Malay Peninsula. *Nusantara* 3:69-79.

KÄHLER, H.

1965 Dialect and language (investigated with some examples from Indonesian languages). *Lingua* 15:497-514.

KASEVICH, V.B.

1983 *Fonologicheskie problemy obshchego i vostochnogo jazykoznanija*. Moscow: Nauka.

KERN, H.

1871 *Kawi-studiën: Arjuna-wiwāha*, zang I en II. The Hague.

NOTHOFER, B.

1975 *The reconstruction of Proto-Malayo-Javanic*. The Hague: Nijhoff.

OGLOBLIN, A.K.

1983a O jazykovyh kontaktah v javanskom areale. *Geneticheskie, areal'nye i tipologicheskie sviazi jazykov Azii*, 115-129. Moscow: Nauka.

1983b Malajskij jazyk i jego dialecty. *Vestnik Leningradskogo universiteta* 2:81-85.

PAVLENKO, A.P.

1965 *Sundanskij jazyk*. Moscow: Nauka.

POERWADARMINTA, W.J.S.

1931 *Baoesastra Djawa I*. Jogjakarta.

RAS, J.J.

1968 Lange consonanten in enige Indonesische talen. *BijdrTLV* 124:521-541.

ROBINS, R.H.

1957 Vowel nasality in Sundanese. *Studies in linguistic analysis*, 87-103. Oxford.

SARKAR, B.

1972 *Corpus of the inscriptions of Java. II*. Calcutta.

SATJADIBRATA, R.

1950 *Kamoes Soenda-Indonesia*. Djakarta.

SHARADZENIDZE, T.S.

1982 *Tipologija jazykov v sinhronicheskom i diahronicheskom plane*. Tbilisi.

SHCHUKO, Ye.M.

1973 Slogovaja struktura dvuhslozhnoj kornevoj morfemy v indonezijskom jazyke. *Sbornik statej po vostochnomu jazykoznaniju*, 174-185. Moscow: Nauka.

SIRK, Ü.

1978 Problems of high-level subgrouping in Austronesian. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, 255-273. PL, C-61.

SNOUCK-HURGRONJE, C.

1892 *Studiën over Atjehsche klank- en schriftleer.* Batavia.

STEVENS, A.M.

1968 *Madurese phonology and morphology.* New Haven.

STOCKHOF, W.A.L.

1980 *Tatabunyi bahasa Indonesia.* *Dewan Bahasa* 24/1:38-54.

TESELKIN, A.S.

1963 *Drevnejavanskij jazyk.* Moscow: Nauka.

TOORN, J.L. van der

1891 *Minangkabausch-Maleisch-Nederlandsch woordenboek.* The Hague.

UHLENBECK, E.M.

1949 *De structuur van het Javaanse morpheem.* Bandung.

WOLF, Ye.M.

1970 *Nekotorye osobennosti struktury slova v portugal'skom jazyke (v sravnenii s ispanskim). Morfologicheskaja struktura slova v indoevropskih jazykah,* 303-346. Moscow: Nauka.

WURM, S.A. and Shirô HATTORI, eds

1983 *Language atlas of the Pacific area, part 2: Japan area, Taiwan, Philippines, mainland and insular south-east Asia.* Canberra: The Australian Academy of the Humanities, in collaboration with the Japan Academy. PL, C-67.

YAKHONTOV, S.Ye.

1971 *Obshchie tendencii razvitija siniticheskikh jazykov. Strany i narody Vostoka* 11:268-279. Moscow.

ZOETMULDER, P.J.

1982 *Old Javanese-English dictionary. I-II.* The Hague: Nijhoff.

ZUBKOVA, L.G.

1970 *O garmonii glasnyh v indonezijskom jazyke. Jazyki Jugo-Vostochnoj Azii: Voprosy morfologii, fonetiki i fonologii,* 216-231. Moscow.

1977 *Konsonantnaja i slogovaja struktura slova v indonezijskom jazyke. Malajsko-indonezijskie issledovanija,* 191-217. Moscow: Nauka.

EASTERN SERAM: A SUBGROUPING ARGUMENT

James T. Collins

0. INTRODUCTION

In a recent report on language research in the Moluccas (Collins 1982b), a proposal for subgrouping the languages of eastern Seram was sketched in an outline form. Within the broad scope of that survey, only a few facts could be noted; so on this occasion I would like to flesh out that outline with more data and a tighter analysis.¹ The task is a very straightforward one: namely subgrouping closely related but hitherto poorly documented languages. The results sharply contrast with the atomistic theories proposed by earlier authors who attempted detailed subgrouping of these languages. Rather, this recent research supports the overview that the Moluccan languages are all interrelated, a position taken in different forms both by Dyen (1965) and Blust (1982 and elsewhere).

If we refer to the map below, it is apparent that, even by East Indonesian standards, East Seram is remote, far from even the modest bustle of Ambon, the provincial capital, and not even near the more frequented parts of West New Guinea.

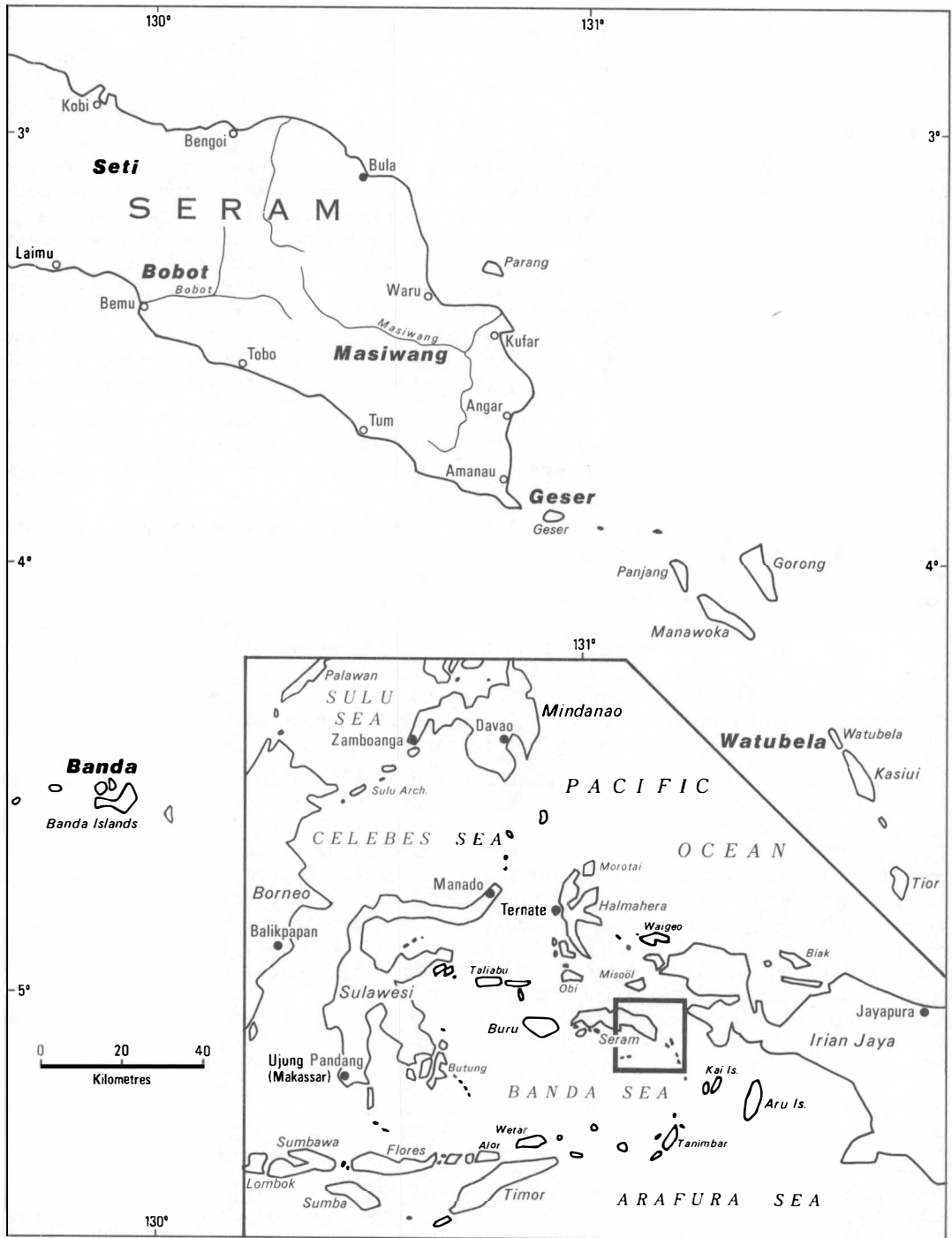
It is no wonder then that very little linguistic research has ever been undertaken there. In this century, only two other writers have attempted language subgrouping in this area. E. Stresemann, working with some 19th century wordlists (Ludeking 1868, Wallace 1869, Boot 1893) and vocabularies collected before 1916 by his colleagues K. Deninger and O. Tauern,² dealt with East Seram in his major attempt at Moluccan language classification (1927). About 50 years later, Chlenov (1969, 1976), drawing chiefly on the same materials used by Stresemann as well as one additional wordlist collected by himself and his wife during their 1960s fieldwork,³ proposed a classification of these languages which differs from Stresemann's in many significant respects.

It is somewhat surprising that these two authors using essentially the same material should have reached conclusions which are quite different. Of course, there is a difference in approach. Stresemann established a taxonomy based on a few linguistic characteristics, while Chlenov's classification is based on lexicostatistical computations. Still, the major factor underlying the disagreement is the lack of information at their disposal. The very paucity and unreliability of the data must necessarily yield multiple, tenuous interpretations. The subgrouping presented in this paper however is based on a new, far greater data base. The six languages discussed here are represented by vocabularies and recordings collected by a single fieldworker over a period of two years.⁴ The data are drawn from, in one case, all the known dialects of the language in question,⁵ and

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 123-146. *Pacific Linguistics*, C-94, 1986.

© James T. Collins

123



Map 1: Eastern Seram

in the other cases from at least three or four dialects of each language. It is hoped that this more representative data base will yield a more reliable subgrouping theory.

The six languages are: Seti, Bobot, Masiwang, Geser, Watubela and Banda. Earlier authors assigned these languages a number of different names and were equally vague about their geographical distribution. Consequently a few remarks about the location and names of these languages are in order (refer to Map).

Seti is a language of the interior of Seram, east of the Isai mountain range and north of the Bobot and Masiwang rivers. Most Seti-speaking villages are now established on either coast of Seram, the westernmost being Seti itself now located on the north coast as one of Wahai's ancillary villages. These Seti-speaking villages include Liambata, Kobi and Benggoi mentioned by other authors.

The Bobot language is spoken in villages along the Bobot river, at its mouth and all along the southern coast of Seram roughly parallel to the river. Stresemann refers to one of these coastal villages, Hatumeten, while Chlenov deals with another, Batuasa. These are both dialects of Bobot.

Masiwang was, until recently, spoken along the length of the great Masiwang river but now it is mostly confined to a string of hamlets north-west of the river's estuary. There is also a very recent resettlement on the south coast of Seram near the Tum river. This language, formerly called Bonfia, has until now only been known through a wordlist collected 90 years ago by a Catholic missionary (Cocq d'Armandville 1901).

Geser, the far-flung language of traders, slavers, sailors and fishermen, is spoken in all but the southernmost Seran Laut islands, all along the north and south coasts of Seram as far west as Hoti and Siloham in the north and is reportedly even spoken in some villages of south-west New Guinea (R. Walker, personal communication). Stresemann referred to this largely as "Seran Laut".⁶ Closely related to it is the Watubela language spoken on Watubela island in the southern part of the Seran Laut group.⁷

The Banda language is no longer spoken in its homeland islands; refugees from the genocide of 1621 fled to Kei Besar in the Kei islands, 400 km to the south-east (see Map), where they established the two surviving Banda-speaking villages, Eli and Elat.⁸ Based on a 19th century wordlist (Van Eijbergen 1865) Stresemann discussed this language as "Eli-Elat".

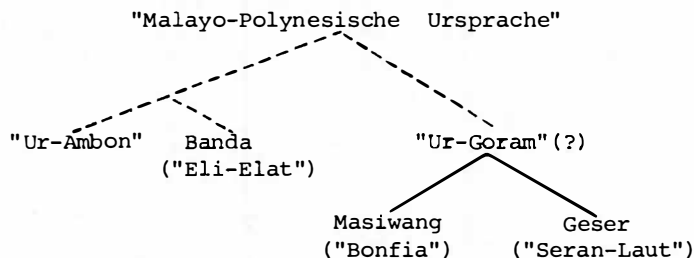
By obtaining a rough geographic notion of range and proximity of these six languages it is possible to understand their interrelatedness. Having identified the languages discussed by Stresemann and Chlenov and established a joint nomenclature for them, it will be easier to study the theories presented by them and to compare those contrasting theories with the proposals set forth here.

1. THE PROBLEM

In the first modern attempt to classify the languages of Central Maluku, Stresemann (1927) proposed a subgroup which he called "Ur-Ambon". He included in this subgroup all the indigenous languages of Buru and Seram and the adjacent islands except Masiwang in East Seram, Geser and Watubela in the Seran Laut islands, and the Banda language. In order to ascertain the boundaries of his Ur-Ambon group, he dealt with these 'peripheral' languages in appendices to his book, sketching sound correspondences of each with Ur-Ambon; he suggests that

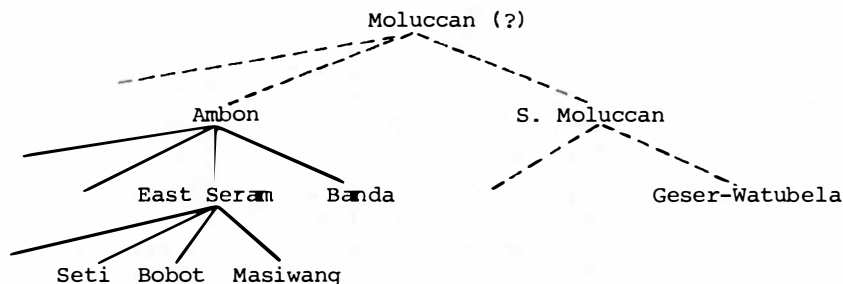
Masiwang and Geser are closely related to each other; indeed he proposes an "Ur-Goram" subgroup from which both are descended. About Ur-Goram he says that it is closer to Proto-Austronesian ("Malayo-Polynesische Ursprache") than Ur-Ambon. The conservative nature of Ur-Goram constitutes the most important difference between it and Ur-Ambon. In contrast, he writes that the Banda language is the language most closely related to Ur-Ambon because there exists a great similarity between the two which indicates earlier common developments.

If we summarise Stresemann's conclusions we reconstruct a family tree of the following shape:⁹



Note that Stresemann emphatically includes Seti ("Liambata, Kobi, Benggoi") and Bobot ("Hatumeten") in "Ur-Ambon", just as he explicitly excludes Masiwang and Geser.

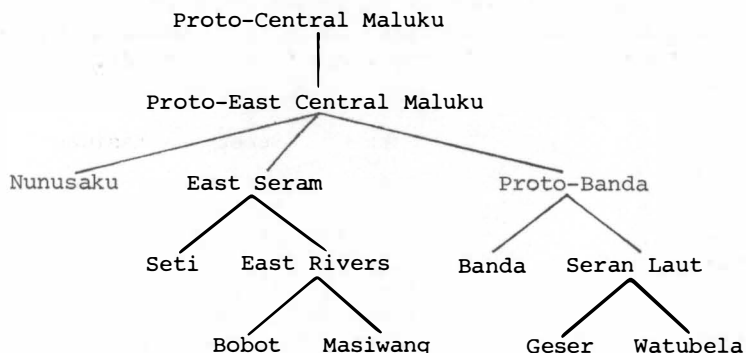
Chlenov (1969, 1976), in sharp contrast to Stresemann, includes Masiwang as well as Bobot and Seti, in the East Seram subgroups of his "Ambon" group. Banda forms a separate subgroup within the same Ambon group. Like Stresemann he assigns Geser and Watubela to the Geser group which he says belongs to the South Moluccan subfamily. This yields a diagram such as:¹⁰



A comparison of the respective theories of Chlenov and Stresemann shows their agreement in placing Geser in a group only related to the languages of Seram ("Ambon", "Ur-Ambon") in a distant, unspecified way. Similarly there is a parallel between the two in assigning Banda a close relationship to the languages of Seram. The most striking point of disagreement is the allocation of Masiwang, Stresemann grouping it with Geser but Chlenov insisting on its close relationship to the other East Seram languages.

In this paper, comparative evidence is offered in strong support of Chlenov's allocation of Masiwang to a group of East Seram languages. The details of the interrelationship of Masiwang, Bobot and Seti with each other are presented as well. However, exception is taken to both Chlenov's and Stresemann's exclusion

of Geser from a Central Maluku subgroup. Indeed, there is evidence that Banda and Geser are closely related to each other and thus sharply distinguished from the languages of South Maluku. A tentative family tree is outlined below:¹¹



2. PROTO-EAST SERAM

Three languages of easternmost Seram display a number of shared innovations which distinguish them both from the languages to the west as well as from those to the south-east (Collins 1982b). Furthermore, they fail to display a number of the innovations characteristic of the language subgroup of West and Central Seram, elsewhere identified as Nunusaku (Collins 1981). For these reasons, Seti, Bobot and Masiwang are grouped together as a major branch of Proto-East Central Maluku: East Seram.

Like other descendants of Proto-East Central Maluku, the East Seram subgroup displays merger of *d, *D, *z and *Z, truncation of diphthongs ending in front off-glides (*ay>a, *uy>u), coalescence of the diphthong *iw to i as well as all the innovations common to all the languages of Central Maluku (Collins 1981). Unlike the Nunusaku languages, however, Proto-East Seram did not merge *R with the reflex of *l. Note the following examples:¹²

PAN		Seti	Bobot	Masiwang
*Rumaq	<i>house</i>	(ufa ^h i ^h ya)	uma	suma
*daRəq	<i>blood</i>	laha	lawa	lasa
*(zZ)uRi	<i>thorn</i>	luila	-lui-	lusi-n
*waRəj	<i>rope</i>	wehela	wawat	wasat
*tiRəm	<i>oyster</i>	tehena	tilan	-
*bulan	<i>moon</i>	(umala)	vulan	hulan
*təlu	<i>three</i>	tol	tolu	toli
*walu	<i>eight</i>	wal	walu	wali
*bulu	<i>body hair</i>	fula ^h a	-vuli-n	hulhuli-n

Although all three East Seram languages display l as the reflex of *l, reflexes of *R are distinct from *l. This contrasts with all Nunusaku languages. In Masiwang the reflex of *R is s (perhaps through devoicing of [ɣ] to [x] and then [s]). In Bobot, with some exceptions,¹³ *R became w (again via [ɣ], presumably); this w was lost when preceded or followed by [u]. In Seti, the correspondence is more problematic. But it appears that *R underwent a conditioned split. Before high vowels *R>l (*(zZ)uRi>luila, *Rusuk>lusu, etc.) but before non-high

vowels *R>h (*daRəq>laha, *waRəj>wehela,¹⁴ etc.). The conclusion is that *R was retained in Proto-East Seram as a sound ([ɣ], perhaps) distinct from *l whereas in Nunusaku *l and *R merged.

The Seti, Bobot and Masiwang languages, which do not undergo this merger of *l and *R display a number of shared innovations which distinguish them from the languages of Nunusaku. First, **ɖ (from *d/D and *z/Z), *l and *j merge to l. For example:

PAN		Seti	Bobot	Masiwang
*qa(zZ)ay	<i>jaw</i>	alam	y-alan	y-alan
*Rəzan ¹⁵	<i>stairway</i>	olam	y-olan	w-olan
*Zalan	<i>road</i>	lala-m	lolan	lolon
*DuSa	<i>two</i>	lua	lua	lua
*daRəq	<i>blood</i>	laha	lawa	lasa
*daqun	<i>leaf</i>	lanna	ai-lan	lan
*bulu	<i>body hair</i>	fula?a	-vuli-n	hulhuli-n
*bulan	<i>moon</i>	(umala)	vulan	hulan
*lima	<i>five</i>	lima	lima	lima
*pija	<i>how many</i>	hila	fila	hila
*Suaji	<i>younger sibling</i>	(aifa)	walin	wali-n
*ŋajan	<i>name</i>	nalan-a	-nalan	nalan

In Collins 1983, the retention of *j and **ɖ as distinct reflexes in Nunusaku is demonstrated; in that branch *l merged with *R.

A second innovation shared by the descendants of Proto-East Seram is the treatment of penultimate *ə. In all three languages, *ə>o/_CV(C)#

*təlu	<i>three</i>	tol	tol	toli
*pənuq	<i>full</i>	bonu	bonu	bom
*dəŋəR	<i>hear</i>	(bulai)	a-vloka	loka
*kəDəŋ	<i>stand</i>	kolo	a-kolan	kolan
*Rəzan	<i>stairway</i>	olam	y-olan	w-olan
*qatəluR	<i>egg</i>	tolla	toli-n	toli-n
*baqəRu	<i>new</i>	folla	voi	hosi
*qapəju	<i>gall</i>	(hoe?a)	foli	holi

The Nunusaku languages (Collins 1983) are reconstructed with a retention of *ə as ə in this position.

Third, while Nunusaku is reconstructed with k as the reflex of *k in initial and medial position, Proto-East Seram is reconstructed with zero as the reflex of *k in those positions on the basis of evidence such as:

*kaSiw	<i>wood</i>	ai-a	ai	ai
*kasaw	<i>rafter</i>	asa-?a	asa-	asa
*kami	<i>we(II)</i>	ami	am	im-
*kulit ¹⁶	<i>skin</i>	inta	kurit	uhit
*takut	<i>fear</i>	mu-ta	m-tait	ka-toit
*ikan	<i>fish</i>	en-a	ian	ian
*sakay	<i>ascend</i>	-	sa:	sa:
*(t)a(ŋ)kaw	<i>steal</i>	-	ma-na:	(kah leit)

Although the three descendants of Proto-East Seram do not agree on the phonetic shape of the reflex of ****nd**, **d** is tentatively reconstructed, based on the occurrence of [d] in Bobot. Masiwang and Seti display **r** but this **r** must have occurred after the shift of ***R** to **s** and **h/l** respectively, as discussed above.¹⁷

In sharp contrast to the languages of Nunusaku, Proto-East Seram must be reconstructed with a retention of ***ŋ** in medial positions. Both Bobot and Masiwang display [k] as the reflex of ***ŋ**. In Seti the reconstruction is a problem because so few cognates containing ***ŋ** occur. Two display **n** while one shows **k**. The partial (?) shift of ***ŋ** to **n** must have occurred after Seti split from Bobot-Masiwang. For example:

PAN		Seti	Bobot	Masiwang
*dəŋəR	<i>hear</i>	(bulai)	a-vloka	loka
*taŋis	<i>weep</i>	tatan	takit	rakit
*laŋit	<i>sky</i>	(lea)	lakit	lakit
*naŋuy	<i>swim</i>	nak	naky	(ahás)
*taliŋa	<i>ear</i>	tinam	talika-n	talikan
*aŋin	<i>wind</i>	(simala)	yakin	yakin
*laŋau	<i>k.o. fly</i>	-	laka	-

Bobot-Masiwang is further distinguished from Seti in the treatment of ***ə** in the final syllable. Note:

*kəDəŋ	<i>stand</i>	kolo	kolan	kolan
*(ma)-qitəm	<i>black</i>	metena	mametan	metan
*waRəj	<i>rope</i>	wehela	wawat	wasat
*qatəp	<i>thatch</i>	ata lanna	yata	-
*daRəq	<i>blood</i>	laha	lawa	lasa
*Daləm	<i>interior</i>	fallal	lala	lalan
*dəŋəR	<i>hear</i>	(bulai)	a-vloka	loka

In Seti, ***ə** assimilates to the vowel of the preceding syllable. In Bobot-Masiwang ***ə > a/ __ (C) #**.

Another characteristic which distinguishes the descendants of Proto-East Seram is the treatment of the ancient noun marker ***si** (Collins 1982a). All three languages display its retention as **y-** in certain words. For example:

*aku	<i>I</i>	ya-ʔa	ya	ya-ʔaho
*Sasaŋ	<i>gill</i>	in-yasan	yasan	yasan
**kambat ¹⁸	<i>wound</i>	il-yabat	yabat	yabat

However in many other words only Bobot and Masiwang display **y-**.

*qa(z)ay	<i>jaw</i>	ala-m	yalan	yalan
*qatəp	<i>thatch</i>	ata lanna	yata	(ilan)
*ama	<i>father</i>	ama-i	yama-n	(baba)
*aŋin	<i>wind</i>	(simala)	yakin	yakin
*asu	<i>dog</i>	(inawa)	yasy	yai ¹⁹

It is possible that these initial **ys** were lost in Seti but retained in Bobot and Masiwang.

An alternative analysis would propose that in Proto-East Seram **y-** (from ***si**) was affixed only to certain words, such as ***aku**, ***Sasaŋ**, etc. After the split of

Seti and East Rivers, the East Rivers branch attached the affix to other words with initial a but Seti did not. At a still later stage, after Masiwang and Bobot had split, Bobot innovated further by affixing y- to other words which had previously been marked with w-²⁰ or zero. Note:

*a <i>q</i> i	<i>leg</i>	wai-na	yai-n	wai
*(R)əzan	<i>stairway</i>	ola	yolan	wolan
*q <i>i</i> Su ²¹	<i>shark</i>	woi ²²	yoi	woi

With regard to the presence or absence of y- it is difficult to determine with certainty which is the retention and which the innovation. But the pattern indicated by these petrified morphosyntactic markers is clear: Seti must have split from Bobot and Masiwang.²³ These data only confirm other evidence of this split (i.e. treatment of *ŋ, *ə, etc.). Some of the differences that distinguish Bobot and Masiwang have been mentioned above. We repeat here two of the most important:

*R	> Bobot	w
	> Masiwang	s
**nd>*dl	> Bobot	d
	> Masiwang	r

Another difference occurs in the treatment of final *u. In Masiwang *u fronts to i; subsequent metathesis occurs when the originally preceding consonant is [n]. In Bobot, however, *u is devoiced or lost (usually after voiced continuants).²⁴ For example:

PAN		Bobot	Masiwang
*batu	<i>stone</i>	vaty	hati
*ma-putiq	<i>white</i>	babuty	buti
*təlu	<i>three</i>	tol	toli
*TukTuk	<i>pound</i>	duty	tuti
*bunuq	<i>kill</i>	vun	huin
*tunu	<i>burn</i>	dun	tuin

Based on materials presented in this section, then, the relationship of the descendants of East Seram to each other and to Proto-East Central Maluku may be represented as follows:

PAN		Banda	Geser	Watubela
*Rumaq	<i>house</i>	rumo	ruma	lumak
*daRəq	<i>blood</i>	raro-	rara	lalak
*(zZ)uRi	<i>thorn</i>	riri-	ruri	luli
*waRəj	<i>rope</i>	warot	(tali)	(tali)
*qaZəy	<i>jaw</i>	ara-	ar	(kakelan)
*Rəzan	<i>stairway</i>	eren	roran	lalon
*DuSa	<i>two</i>	ruo	ro-ti	lua
*daqun	<i>leaf</i>	ran-o	ru	lue?
*pija	<i>how many</i>	ilo	fis	fihi
*ŋajan	<i>name</i>	nalan	ŋasan	ŋahan
*bulan	<i>moon</i>	ɸulan	ulan	ulan
*lima	<i>five</i>	limo	lim	lima

In both Banda and Geser the reflex of *d/D, *z/Z and *R is [r]; this r shifted later to [l] in Watubela. The merger of *d/D, *z/Z and *R is, thus, considered an innovation shared by all the descendants of Proto-Banda. Note that the reflex of *j is distinct from this r in all three descendant languages. Geser displays [s] as the reflex of *j (thus, merging with *s); Watubela has shifted all s's to [h]. As noted above Watubela has merged an earlier r (from *d/D, *z/Z and *R) with *l. In Banda *j merges with *l as l. Geser alone retains the original three-way distinction; that is:

*d/D, *z/Z, *R > r
 *s, *j > s
 *l > l

It is precisely this joint treatment of this series of voiced apicals which justifies the positing of a Proto-Banda subgroup. Note that in Nunusaku *R and *l merged but *d and *j did not; while in Proto-East Seram it was **ɖ, *j and *l which merged but not *R. Although the contemporary reflexes vary (due to later shifts and mergers), the pattern displayed in Banda, Geser and Watubela is the same.

Numerous innovations, however, distinguish Banda from the Seran Laut branch, consisting of Geser and Watubela. First, we mention again the shift of *j to l in Banda which contrasts with the merger of *j and *s in Seran Laut. Second, the merger of *ŋ with *n in Banda contrasts with the retention of *ŋ as [ŋ] in initial and medial positions in both Watubela and Geser. For example:

PAN		Banda	Geser	Watubela
*naŋuy	<i>swim</i>	nano ²⁶	naŋu	(gaha)
*taŋis	<i>weep</i>	(ndaot)	taŋis	ntaŋi
*təliŋa	<i>ear</i>	tilu-	tiliŋa	telŋa
*dəŋəR	<i>hear</i>	(motan)	roŋan	doŋan
*saŋa	<i>branch</i>	sana-	saŋ	(lifā)
*aŋin	<i>wind</i>	anin	aŋin	aŋin
*ŋajan	<i>name</i>	nalan	ŋasan	ŋahan

With only one exception (laŋit, an apparent loanword, perhaps from Malay), *ŋ > n in Banda.²⁷ In all cases Geser and Watubela retain [ŋ].

Third, in Banda *ə became [o] in final position but [e] in penultimate position. The few exceptions seem to be conditioned by assimilation. In Geser and Watubela the split is between [a] in final position and [o] in the penultimate syllable. For example:

PAN		Banda	Geser	Watubela
*qatəp	<i>thatch</i>	atop	(barem)	kataf
*daRəq	<i>blood</i>	raro-	rara	lalak
*waRəj	<i>rope</i>	warot	(tali)	(tali)
*ma+qitəm	<i>black</i>	metmetén ²⁸	metan	maketan
*Daləm	<i>interior</i>	raron	lomin ²⁹	lomi
*ənəm	<i>six</i>	nemu ³⁰	onan	onon
*dəŋəR	<i>hear</i>	(motan)	ronan	donan
*təlu	<i>three</i>	telu	tolu	tolu
*pənuq	<i>full</i>	mbunu	(lome)	(lomi-n)
*ma+qəti	<i>ebb</i>	meti	moti	mkoti
*qatəluR	<i>egg</i>	tuluru	tolu	katlu
*qapəju	<i>gall</i>	elu-n	foli	feli-n
*baqəRu	<i>new</i>	ferferu-ŋo	wou-wou	oku-oku

Fourth, the treatment of the labial stops distinguishes Banda from the descendants of Seran Laut. Note:

PAN		Banda	Geser	Watubela
*qapəju	<i>gall</i>	elu-n	foli	feli-n
*paRi	<i>rayfish</i>	ari	fari	fali
*xapuy	<i>fire</i>	au	afi	afi
*qatəp	<i>thatch</i>	atop	(barem)	kataf
*bəRsay	<i>paddle</i>	fese	wosa	woho
*batu	<i>stone</i>	fatu	watu	watu
*busuR	<i>bow</i>	ɸusur	usu	uhul
*bubu	<i>trap</i>	ɸuɸu	u:	u:
*əmpu	<i>grandparent</i>	umbo	(tata)	(tata)
*(ma)+putiq	<i>white</i>	noito-no	futi	buti? (<i>skin fungus</i>)
**kambat	<i>wound</i>	ambat	abat	abat

These data can be summarised as follows:

- *p > Banda \emptyset / { $\frac{\#}{\underline{v}} \underline{v}$ }, p/___#
 > Seran Laut f
- *b > Banda f/___ $\begin{bmatrix} +voc \\ -rnd \end{bmatrix}$, ɸ/___ $\begin{bmatrix} +voc \\ +rnd \end{bmatrix}$
 > Seran Laut w/___ $\begin{bmatrix} +voc \\ -hi \end{bmatrix}$, ɸ/___ $\begin{bmatrix} +voc \\ +bk \\ +hi \end{bmatrix}$
- *mb, *mp > Banda mb
 > Seran Laut b

Fifth, the treatment of the reflexes of prenasalised labials noted above is paralleled by the treatment of prenasalised apicals. For example:

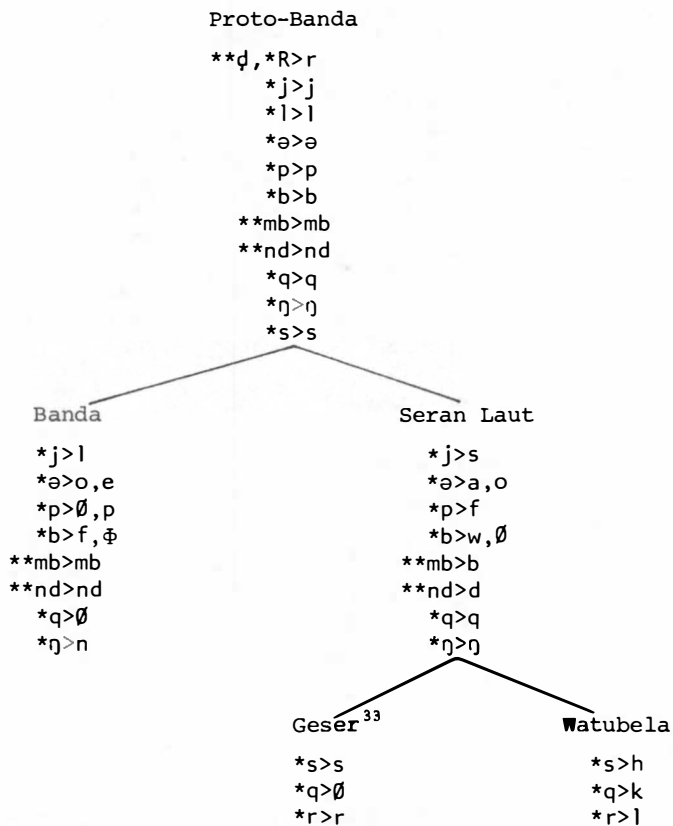
PAN		Banda	Geser	Watubela
*punti	<i>banana</i>	(kula)	fudi	fudi
*ma+ttunu	<i>burn</i>	ndunu	(tunu)	dunu
*diŋdiŋ	<i>cold</i>	rindin	(rifi)	lidin
*DaRat	<i>land</i>	ndara-k ³¹	(nodi)	(na-kodi)

In Banda *nt and *nd merged as the cluster [nd] but in Seran Laut **nd (from *nt and *nd) has become [d].

Sixth, Watubela preserves a distinct reflex of *q in all positions.³² Although Geser has lost *q, we reconstruct it in Seran Laut but not in Banda. Note:

PAN		Banda	Geser	Watubela
*qatəluR	<i>egg</i>	tuluru	tolu	katlu
*qatəp	<i>thatch</i>	atop	(barém)	kataf
*ma+qəti	<i>ebb</i>	meti	moti	mkoti
*baqəRu	<i>new</i>	ferferu-	wou-wou	oku-oku
*Rumaq	<i>house</i>	rumo	ruma	lumak
*daRəq	<i>blood</i>	raro	rara	lalak

Based on the evidence presented here we reconstruct a family tree for Proto-Banda below:



4. CONCLUDING REMARKS

This paper proposes a classification of six languages of eastern Seram and the adjacent islands. The subgrouping is based on shared phonological innovations, a traditional, if beleaguered, approach. The analysis proposed here divides the six languages into two groups, East Seram and Proto-Banda, both of which are sister languages descended from Proto-East Central Maluku. The inter-relationships of the languages within each of these subgroups has been worked out and tentative family trees have been proposed.

Clearly, this classification is an initial step in the long process of refining knowledge and theories about these languages. Not only do we need more information about the languages but furthermore we need more comprehensive methodological approaches which rest on morphological as well as phonological innovations and which are more sharply aware of the sociocultural factors, indeed the sociolinguistic history of the area.

Nonetheless, even this rudimentary sifting of evidence suggests a theory of classification which can be challenged and tested. It is apparent now that the chief differences between most of these languages and those further to the west occur because the eastern Seram languages are characterised by retentions of sounds and sound patterns which have been largely lost in western Seram. The identification of these retentions, *q, *ŋ, *j, *ə, *mb/*nd, is important in two ways.

First, it provides a picture of Proto-Central Maluku's phonological system which is strikingly different from the one proposed in Stresemann 1927. It is precisely these differences from Stresemann's model which yield a Proto-Central Maluku phonological system which is richer, more complete and much more similar to the system proposed for Proto-Austronesian. Proto-Central Maluku, then, can be considered much more conservative than a study of only western Seram languages would indicate.

Second, in Collins 1983 certain hypotheses were set forth regarding the reconstruction of Proto-Central Maluku *q, *j and *ə. These theories were based on sound correspondences in some languages of western Seram and the somewhat complex arguments needed to interpret them. The data set forth in the present paper provide strong support for those hypotheses and justify the interpretation proposed there.

Certainly some aspects of the subgrouping theory presented in this paper are stronger than others. The close relationships of Bobot to Masiwang and of Geser to Watubela seem sufficiently clear. But, for example, the relationship of Seti to Bobot and Masiwang requires closer scrutiny. A more detailed comparison of Seti to the languages immediately westward, in particular variants of Manusela, is an obvious project for future research. The collection of a larger body of information regarding Seti dialects is also imperative. The relationship of some languages, for example Salas, to other variants of Seti is none too clear, largely because the amount and quality of the data collected are far from satisfactory.

Indeed, perhaps this attempt at subgrouping the languages of eastern Seram has been premature. One would have preferred that detailed phonetic sketches, grammatical outlines and extensive vocabularies for each of the languages compared were prepared first, before subgrouping proceeded. Attempting classification on the basis of a few hundred or even a few thousand words is not a model approach for comparative linguistic research. So, this paper is presented as a starting point, a testable theory, a flawed beginning. But a beginning has been made.

NOTES

1. The research upon which this and earlier papers are based was conducted under the auspices of Lembaga Ilmu Pengetahuan Indonesia; the support and advice offered by the staff there is gratefully acknowledged here. The provincial government of Maluku and its officers fully cooperated in my efforts to collect data. Furthermore, I owe a special word of thanks to F. Berhиту and his family who were my hosts in Geser and to A. and L. Rumra of Elat who assisted me in collecting information about the Banda language, now spoken on Kei Besar. But my research would not have been possible without the generosity and hospitality of the many people in Maluku who helped, tolerated and befriended me. A partial list of informants is included in Collins 1983. My thanks also go to Sergei Ignashev for his assistance with the Russian materials and several obscure bibliographical references.
2. Stresemann (1927:2) acknowledges that Deninger collected the wordlist for Hatumeten, which was his chief source for the language named Bobot in the present paper. Tauern also sent him information about Liambata and other East Seram languages.
3. Chlenov (1976:226) reports that he recorded a 507 word vocabulary and 36 sentences of a variant of Gorom (called Geser here). The material was collected from a native of АМАРБАВАТУ, Amarwawatu (?) but apparently Amarwatu on the south-easternmost cape of Pulau Gorom Laut. In addition to the materials at Stresemann's disposal, Chlenov also referred to several other wordlists including Miklucho-Maclay 1950, 1951 (based on that Ukrainian ethnologist's visit to south-east Seram in 1874; see Cowan 1957: 290), Rosenberg 1878, Riedel 1886 and Ribbe 1892.
4. The data collection took place during my doctoral field work, September 1977 to August 1979.
5. The Banda language apparently survived in only two variants, Eli and Elat, on Pulau Kei Besar. See note 8 for an explanation.
6. Stresemann's information on this language group seems to have been drawn from Cocq d'Armandville 1901 and Riedel 1886, but S. Müller and N.N. Miklucho-Maclay are mentioned as well (Stresemann 1927:199).
7. Riedel's (1886) information, though scattered and sparse, is fairly reliable, but Wallace's "Matabello" wordlist (1869) appears to be from a less conservative dialect of Kesui Island, just south-east of Watubela Island proper. It seems that Wallace did visit Muslim Watubela (which he called "Kisiwoi" (1869:279)) but the wordlist he published was collected from pagans on Kesui (which he called "Uta").
8. For an account in English about the murderous depopulation of the Banda Islands see Hanna 1978. (Apparently Coen, the Governor General, was simply following a decision made as early as 1615 in Holland by the directors of the VOC; refer to Hyma 1953:11.) Their flight and resettlement on Kei Besar was well known (Van Eijbergen 1865); indeed it is reported as early as 1656 (Leupe 1875).
9. The details of the relationships among Ur-Ambon, Eli-Elat, Ur-Goram and PAN ("Malayo-Polynesian Ursprache") are not clear.

10. Note that the dendrograms summarising the conclusions of both Chlenov and Stresemann were designed by this writer. Every attempt has been made to avoid error and distortion but of course it may be misleading to diagram verbal explanations, especially those based on lexicostatistical computations.
11. Some of the details of the subgrouping theory implied here are presented in Collins 1981, 1982b, 1983.
12. Reconstructions used here are largely drawn from Wurm and Wilson 1975; however, PAN *e, a central vowel, is written ə to avoid confusion with mid-front vowels in some languages of Maluku.
13. Many of these exceptions are the names of sea creatures, for example, *oyster* in the preceding list. Because the Bobot peoples were probably originally people of the interior (along the Bobot River), these sea creature names may be loanwords.
14. There are other irregularities in these two examples.
15. The abbreviated reconstruction *Rəzan, rather than *SaRəzan appears here.
16. The Seti reflex displays metathesis followed by syncope and subsequent assimilation; that is **ulita > iluta > ilta > inta. The Bobot reflex unexpectedly displays a k.
17. There are at least two other minor innovations which the three descendants of Proto-East Seram appear to share. After the loss of *q, petrified prefix *ma- before i in *ma-qitem *black* yielded me-, note:

PAN	Seti	Bobot	Masiwang
*ma+qitem	metena	ma-metan	metan

Furthermore, in the numeral series, the reflex of *ənəm in all three languages displays o: Seti noi, Bobot nom, Masiwang nom. In Seti, this would be the expected reflex but in the others it is not: a shared irregularity?

18. This is not a Proto-Austronesian reconstruction but, as the double asterisk indicates, a lexeme tentatively reconstructed for Proto-Central Maluku.
19. The loss of *s is unexplained here.
20. These cases of w- are thought to be reflexes of the Proto-Austronesian article *u (Collins 1982a).
21. The reflex of *qiSu in the Central Maluku languages will be treated in a forthcoming note.
22. This lexeme was not collected in Seti village but in Salas. In six other dialects of Seti, the form varies. On the south coast, Seti-speaking villages display forms influenced by neighbouring languages. Uluhanan, for example, has wokan, clearly a borrowing from the East Littoral languages which border it. In Lesa and Adabai, both within easy walking distance from Atiahu, where a dialectal variant of Bobot is spoken, forms which are similar to Bobot occur (yoyam, yoya respectively). On the north coast, Benggoi and Kobi (tentatively considered variants of Seti) display wuyam and woia. Salas (probably a Seti dialect) shows woï but here, too, proximity to Masiwang-speaking villages may be a factor.
23. Note, for example, the distribution of Seti waha *fire* compared to Bobot yafʋ and Masiwang yah.

24. Note that in Bobot *vovoi new* is an apparent exception. Furthermore, in the Atiahu variant of Bobot -u is lost except when following [h]. Note the following comparisons between Atiahu and Werinama on opposite banks of the Bobot (though each is some kilometres away from the river).

PAN	Atiahu	Werinama
*batu <i>stone</i>	hat	vat
*təbu <i>sugarcane</i>	tohon	tov
*bubu <i>fish trap</i>	huhu	vuv

Apparently then Bobot (if we are to account for the Atiahu reflexes) retained final [u] when preceded by [β]. After [β] became [v] in most Bobot dialects that u was lost too. In Atiahu β>h and -u was retained.

25. There do however, appear to be some problems with the reflexes of *uy in Geser.
26. In Banda *nano wade neck deep*, not *swim*.
27. Note the error in Collins 1982b:121. "Banda retains *ŋ as [ŋ]" should read "Geser retains *ŋ as [ŋ]".
28. The stress on the final syllable suggests that a suffixed genitive marker -ni ~ -in has changed the expected final a to e, i.e. *metmetanni* > *metmetenni* > *metmetén*.
29. The addition of -i(n) occurred before the shift of *ə to o, a.
30. The Banda form is an exception.
31. *DaRat *shore, land (not sea)* with a presumed verbal affix *ma, yields *ndara-k go aground* in Banda.
32. There are a few exceptions. Note, for example No.180 of the appended comparative wordlist. One would predict *lauk* as the reflex of *Zauq *far, distance*. In this case, perhaps *lau far* has merged semantically with *lau seaward, across the sea*.
33. Some dialects of Geser display additional innovations. In Pulau Panjang, for example, [f], the reflex of *p, has moved to [h]; *k shifted to [ʔ] whereas it is retained in Geser. In Kiandarát word-final *u shifted to [i] and was metathesised, yielding forms like [wait-a] *stone* < *batu, [suis-a] < *susu *breast*, etc.

REFERENCES

- BLUST, R.
1982 The linguistic value of the Wallace Line. *Bijdragen tot de Taal-, Land- en Volkenkunde* 138:231-250.
- BOOT, J.
1893 Korte schets der Noord-kust van Ceram. *Tijdschrift van het Nederlandsch Aardrijkskundig Genootschap* 10/2:1183-1190.

CHLENOV, M.A.

- 1969 Ocherki po etnicheskoi istorii narodov tsentral'nykh molukk (Indoneziia). Moscow, dissertation.
- 1976 *Naselenie molukkskikh ostrovov*. Moscow: Nauka.

COCQ D'ARMANDVILLE, C.I.F. le

- 1901 Vergelijkende woordenlijst in het Hollandsch, Ceramsch, en Alfoersch (Bonfia). *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 43: 349-396.

COLLINS, J.T.

- 1981 Preliminary notes on Proto-West Central Maluku: Buru, Sula, Taliabo, and Ambelau. In R.A. Blust, ed. *Historical linguistics in Indonesia*, Part I, NUSA 10:31-45. Jakarta: Badan Penyelenggara Seri NUSA.
- 1982a Prothesis in the languages of Central Maluku: an argument from Proto-Austronesian grammar. In Amran Halim, Lois Carrington and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, vol.2: *Tracking the travellers*. PL, C-75:187-200.
- 1982b Linguistic research in Maluku: a report of recent field work. *Oceanic Linguistics* 21:73-146.
- 1983 *The historical relationships of the languages of Central Maluku, Indonesia*. PL, D-47.

COWAN, H.K.J.

- 1957 Korte mededelingen: een Russisch gegeven over de taal van Watubella. *Bijdragen tot de Taal-, Land- en Volkenkunde* 113:290-291.

DYEN, I.

- 1965 *A lexicostatistical classification of the Austronesian languages*. *International Journal of American Linguistics*, Memoir 19. Baltimore.

EIJBERGEN, H.C. van

- 1865 Korte woordenlijst van de taal der Aroe- en Key-Eilanden. *Tijdschrift voor Taal-, Land- en Volkenkunde* 14:557-568.

GOWARD, N. and R. ELLEN

- 1981 A wordlist of the Bati language. MS in the possession of the author.

HANNA, W.A.

- 1978 *Indonesian Banda: colonialism and its aftermath in the Nutmeg Islands*. Philadelphia: The Institute for the Study of Human Issues.

HYMA, A.

- 1953 *The Dutch in the Far East*. Ann Arbor: George Wahr.

LEUPE, P.A.

- 1875 De reizen der Nederlanders naar Nieuw-Guinea en de Papoesche eilanden in de 17e en 18e eeuw. *Bijdragen tot de Taal-, Land- en Volkenkunde* 22:1-311.

LUDEKING, E.W.A.

- 1868 *Schets van de Residentie Amboina*. The Hague: M. Nijhoff.

MIKLUCHO-MACLAY, N.N.

- 1950 Putesestvie iz Amboiny na bereg Papua-Koviaj (dnevnik putesestvija). *Sobranie socinenij*, vol.2. Moscow: Nauka.
1951 Spisok nekotoryx slov dialektov papuasov berega Papua-Koviaj na Novoj Gvinee. *Sobranie socinenij*, vol.3, Pt.1. Moscow: Nauka.

RIBBE, K.

- 1892 Ein Aufenthalt auf Gross-Ceram. *XXII Jahresbericht des Vereins für Erdkunde zu Dresden*.

RIEDEL, J.G.F.

- 1886 *De sluiik- en kroesharige rassen tusschen Selebes en Papua*. The Hague: M. Nijhoff.

ROSENBERG, H. von

- 1878 *Der malayische Archipel*. Leipzig: Gustav Weigel.

STRESEMANN, E.

- 1927 Die Lauterscheinungen in den Ambonischen Sprachen. *Beiheft der Zeitschrift für Eingeborenen-Sprachen*, 10. Berlin: Dietrich Reimer.

TAUERN, O.D.

- 1931 Beitrag zur Kenntnis der Sprachen und Dialecten von Seran. *Anthropos* 26:109-139.

WALKER, R.

- 1978 A wordlist of the Goram language, taken at Kaimana, Irian Jaya. MS. SIL, Jayapura.

WALLACE, A.R.

- 1869 *The Malay Archipelago*. London: Macmillan.

WURM, S.A. and Shirô HATTORI, eds

- 1981 *Language atlas of the Pacific area, Part 1: New Guinea area, Oceania, Australia*. PL, C-66.
1983 *Language atlas of the Pacific area, Part 2: Japan area, Taiwan, Philippines, mainland and insular South-East Asia*. PL, C-67.

WURM, S.A. and B. WILSON

- 1975 *English finderlist of reconstructions in Austronesian languages (post-Brandstetter)*. PL, C-33.

APPENDIX:

A comparative wordlist of the languages of Eastern Seram

In the following pages the six languages discussed in this paper are represented by a standard 200 word list (as revised by R.A. Blust). Because materials were collected for more than one variant of each of the languages, it has been somewhat difficult to organise these lists. In the end, one variant was selected for each language; occasionally material was not available for the selected variant so in those cases the data was supplied from another variant. So unless otherwise stated the data presented below come from the following sources:

Seti:	Seti village
Bobot:	Werinama
Masiwang:	Kota Baru
Geser:	Selor
Watubela:	Efa
Banda:	Elat

With regard to the phonetic notation, stress is penultimate unless otherwise marked. A dash at the beginning or end of a word indicates that it usually appears with person marker. The notation _◌ (written below a vowel) indicates devoicing; y indicates an unspecified devoiced vowel. A semicolon between entries means that more than one word recurs with roughly the same meaning. A slash between two words indicates that these forms appear in morphologically complementary patterns. Parentheses around the initial letter have the same meaning; for example in Seti *kai* and *ai* are morphological variants in an inflected verb system; so we write (k)ai. A plus sign within a word indicates a morphologically complex construction. Some predictable phonetic alternations are not represented in these lists; for example, in Banda /f/ always appears as [ɸ] before round vowels but only *f* is written in the list.

To reduce the number of notes to the Appendix, letters have been written behind a word that is not taken from one of the six dialects mentioned above. The letter initial in parentheses after the word is the first letter of the dialect from which the word has been taken. The codes are as follows: For Seti, (A): Adabai; for Bobot, (N): Naiyaba, (T): Tobo; for Masiwang, (D): Dawang; for Geser, (G): Geser, (K): Kilmuri, (P): Pulau Panjang; for Watubela, (K): Keldor; and for Banda, (E): Eli.

	Seti	Bobot	Masiwang	Geser	Watubela	Banda
1. <i>arm</i>	ima-	-niman	hasan	ma	ma-	lima-
2. <i>left</i>	ibaVa	avet	kabalit	kubalit	balit	fitar
3. <i>right</i>	fiti fita	kamuin	lalait	kuanan	mela	ratu
4. <i>leg</i>	wai-	-yain	wai	ke	kwe-	ai-
5. <i>walk</i>	rak/tak	ko	uha	tangi	fana	mbo
6. <i>path</i>	lalam	lolan	lolon	lalano (K)	-	eren
7. <i>come</i>	forok	slot	boti	ratan	elat	oma
8. <i>str</i>	bui	butar	hutan	futar	futal	mbutor
9. <i>swim</i>	nak	naky	ahás	naŋu	gaha	nano ¹
10. <i>dirty</i>	kailoa (A)	fof ₃ fof ₃	ba?obo?	galotak (K)	muden	moŋjón
11. <i>dust</i>	kaitaha lafna	lav	ubas	ubas	uban	ofin
12. <i>skin</i>	inta	-kurit	uhit	likit	kalikit	kilit
13. <i>back</i>	sis-	-kotan	karan	kotan	kotoni-	miri-
14. <i>belly</i>	tai- ²	-tian	tuan	toa	kabu-	tia-
15. <i>bone</i>	lúil-a	-luin	lusin	ruri	luli	riri
16. <i>intestine</i>	tua-na	-vituan	tuan walan	utuk	utka-	tain kambun
17. <i>liver</i>	lali-;ata ³	-yatan	yatan	at	lamno-	raro-
18. <i>breast</i>	susa	-susin	susin	susu	susu	susu-
19. <i>shoulder</i>	mabala- (A)	-vayayan	malat	wakir	wawa-	fara-
20. <i>know</i>	mane	vaneta	kalelan	ruk	lunam	mbo+kaiko
21. <i>think</i>	bisna	fikir	ita	hi?ir (P)	-	m+fikir
22. <i>fear</i>	muta	mtait	katoit	matakut	matakut	mbo+takunia-
23. <i>blood</i>	láha	-lawa	lasa	rara	lalak	rara-
24. <i>head</i>	ulu- ⁴ (A)	-ulin katin (N)	ulin	lu	kulu-	ulu-
25. <i>neck</i>	solo-	-yonin	wosan	totolan	aliko-	enu-
26. <i>hair</i>	elfua	-ulin	ulin lan	uk	uka	fuk-
27. <i>nose</i>	ninu-	-ilin	ninin	so	hu-	nilu-
28. <i>breathe</i>	bo'o nawa (A)	-mesi	manowa	fai ŋak	fail uwak	mbilik nuwa-
29. <i>sniff</i>	mala	-bufmauk	mulit	fimusi	kuŋahi	mbok fau
30. <i>mouth</i>	furu-	vivin;vudin	uban	lo	hili	suo-
31. <i>tooth</i>	nesi-	-nifa(n)	nihan	ŋisik	nifo	ni-
32. <i>tongue</i>	lela-	-melin	leha?	kel	lama	me-
33. <i>laugh</i>	mamal	mli	balí?	malif	mumlifa	mbo+malik
34. <i>weep</i>	tatán (A)	takit	rakit	tarjís (G)	ntarji	ndaut
35. <i>vomit</i>	momuta	-lua	lua?	luak	luak	luako
36. <i>spit</i>	fi to/bito	beba	batof	ginisuk	mukanihuk	mbifiru
37. <i>eat</i>	(k)ai	ken	kan	ngga faŋa	golat	mban
38. <i>chew</i>	lamtak (A)	mama	-	mama (G)	mamak	mamo
39. <i>cook</i>	bakalai (A)	-masa	hakatuin	masak	mahak	mososak
40. <i>drink</i>	(k)unu	ninŋ	nimin	inu	minu	-inu
41. <i>bite</i>	(k)oto	kat	kasat	karat	gukut	-iki
42. <i>suck</i>	sus;mo	moit;sosa	komoan	bus;sorat	bus;busa	mbuk ⁵
43. <i>ear</i>	tina-	talikan	talikan	tiliŋa	telŋa-	tilu-
44. <i>hear</i>	bulai	bloka	loka	roŋan	doŋan	motan
45. <i>eye</i>	mata	matan	mata	mata	mata	mata-
46. <i>see</i>	(k)ote	bla	ita	mekun	sini?	-ito
47. <i>year</i>	(k)amowo	mamawa	mumáh	maflaba	ma:f	nuak (E)
48. <i>sleep</i>	(k)ina	mtolŋ	nima	kifit	gena	muturu
49. <i>lie down</i>	kemeluhu	burburij	nima kanowa	bula-bula	bula bula	muturu mbolor
50. <i>dream</i>	mihin (A)	mnif	maníh	manifi (G)	-mifi	ndososuo

	Seti	Bobot	Masiwang	Geser	Watubela	Banda
51. <i>sit</i>	huttue/buttue	duwat	tuat	matoran	sabóu	metet
52. <i>stand</i>	(k)olo	kolan	kolan	mariri	mili?	meliri
53. <i>person</i>	mansia	mansia	mansia	mančia	mansia	mapčia
54. <i>man</i>	mualáina	muana	masanan	urana	marjkana	morana
55. <i>woman</i>	hifnáina	vina	baroha	wawina	hilala	meifino
56. <i>child</i>	yana?a	yana	o?on	anak	ananak	anako
57. <i>husband</i>	mualáina	sawan	masanan ⁶	rana	marjkana ⁷	tau- morana
58. <i>wife</i>	fifnáina/sawa- ⁸	sawan	baroha	wawina	hilala	tau- meifino
59. <i>mother</i>	inai	inan	ina	nina	nina	ina- ⁹
60. <i>father</i>	amai	yaman;baba	baba	baba	yai	ama- ⁹
61. <i>house</i>	ufaiYa	uma	suma	ruma	lumak	rumo
62. <i>roof</i>	ata lanna	yata	ilan	barém	kataf	atop
63. <i>name</i>	nalana	nalan	nalan	ņasan	ņahan	nalan
64. <i>say</i>	rindak/dindak	dula	lian	kuk	mamán	lia-
65. <i>rope</i>	ai wehela	wawat	han wasat	tali (G)	tali	warot
66. <i>tie</i>	lahe	nafit	tes	wokas	deta	lakot
67. <i>sew</i>	labe	bakasoma	tula	firait	famlait	mborait
68. <i>needle</i>	lóusa	lalayan (N) (T)	lalain	lalan	lalan	bilin
69. <i>hunt</i>	ila;usir	vakavela	ukikia	leus;musir	lela	rasté
70. <i>shoot(bow)</i>	hana	bana	hana	fanak	fanak	mbano
71. <i>stab</i>	luk	devat	toi?	tewat	tulak	ndusuk
72. <i>strike</i>	fele/bele	biat	lait	sakur	ņugul	ndukul
73. <i>steal</i>	kamana (A)	mana:	kahleft	fakaleus (G)	komanaka	mbo+naka
74. <i>kill</i>	humata/bumata	vunŷ	huin	tara ¹⁰	bunuk	mbo+kota
75. <i>dead</i>	mata (A)	mata	mata	-mata	-mata	mata-
76. <i>alive</i>	ninawa	meka (N)	lihilihan	liu	mliu	nunuli
77. <i>scratch</i>	(k) fala	kwas	kasa	fikori	goli?	mbo+skakir
78. <i>cut</i>	fita/bitá	da:	tasa	lait	mlait	mbutu
79. <i>wood</i>	ai-a	ai	ai	kai	kai	uno
80. <i>split</i>	bolo	bat	bola ha?as	tola	bakas	mbo+fkila
81. <i>sharp</i>	moi	kat	moit	matarin	matarin	mout
82. <i>dull</i>	koko	nona	babukit	baba	-kati	tutal
83. <i>work</i>	-	dei	-	-	garja	mbuno
84. <i>plant</i>	so?ak	bakatáu	toi?	tauk	daun	ndano
85. <i>choose</i>	liike (A)	badali	pili	filik (G)	fie	mbili
86. <i>grow</i>	abot (A)	me	-	-	-	mbotuk
87. <i>swollen</i>	boto	bivawa	boka	bulu	-bulul	ndofaro
88. <i>squeeze</i>	kisa?at	lefa	-	singatan (X)	dindis	ndembet
89. <i>hold</i>	(k)opla	doman	sit	dahan	gudi?	ndakan
90. <i>dig</i>	(k)ali	kalŷ	kais	kayk	galik	mbaéro
91. <i>buy</i>	safe (A)	balat	has	fas	sawal	mbaso
92. <i>open</i>	sire (A)	balak	sia?	salik	badil	mbende
93. <i>pound</i>	dut (A)	dutŷ; dau	tuti	tutu	dutu	ndutu
94. <i>throw</i>	tafa/rafa	dutŷ	toi?	folá	dulak	ndutu
95. <i>fall</i>	buk	mnav	rei?	ņgek	dakadokit	monafu
96. <i>dog</i>	inawa	yasŷ	yai	kafuna	kofuna	asu
97. <i>bird</i>	manua	miova	ohas	manuk	manuk	manuk
98. <i>egg</i>	tolla	tolin	tolin	tolu	katlú	tulur
99. <i>feather</i>	ful?a	vulin (T)	hulhulin	wilfu	walu-	fulu-n
100. <i>wing</i>	ifale-na	yeun	hián	waki	waliunfe	foro

	Seti	Bobot	Masiwang	Geser	Watubela	Banda
101. fly	tif/rif	div	rihi	tiu	dekan	ndifu
102. rat	ulá:fa (A)	milava	ilaha	kalufa (G)	-	folafa
103. flesh	isi	-isin	sama?	sisia	hai	isi-
104. fat	mina (A)	-iak	mina? (D)	wodal ¹¹ (G)	babal ¹¹	loloi-
105. tail	etna	efat	lekan	kaku	sku	keto-
106. snake	ula (A)	bufin	ula	tekis	utolok	nia ¹²
107. worm	sinalo ^{wa}	urat ¹³ (T)	bat kutin	ena kutu	kuwatawatal	tambir luano
108. louse	fatana	kutin (N)	taniman	kutu	kutu	kutu
109. mosquito	uhua (A)	umanis	-	umis (G)	umus	lokumut
110. spider	lawa lawa (A)	-	-	gafas (G)	la:k	lau
111. fish	ena	ian	ian	ikan	ikan	ikan
112. rotten	fáuna (A)	koban	ma?obon	ngofak	-	fau-
113. branch	sanna	ai yakan	saksakan	san	lifa	sana-
114. leaf	lan	ai lan	han lan	ru	lue?	rano
115. root	tap la	ai wa?at	tamit	akar	akal	wakat
116. flower	kof la	ai futin	dulan husin	fus	fuha	saur
117. fruit	fona-	ai vuan	huan	woi	oke?	fuan
118. grass	ai fafuta	fof ^y	hanulin	imak	egegil	furun
119. earth	kaitaha	yena	bata	ena	tana	tambir;nonain
120. stone	fat ^{wa}	vat ^y	hati	watu	watu	fatu
121. sand	enna	yena	bes	ena	kena wo	nui
122. water	waya	wai	way	ar	-	war
123. flow	abán	bait	roi?	tauk	-	ŋJok
124. sea	taisá	tasi	tasi?	tasik	tahit	tasik
125. salt	táisa	-	-	-	-	morosan
126. lake	tor ^{wa}	tif ^y	balabalan	louk ¹⁴	kahar	kual
127. forest	ai dimana	ai lalan	latan	imak;liba	ehi lome	mur wasi
128. sky	leá	lakit	lakit	lanjit	lanjit	lanjit
129. moon	umala	vulan	hulan	ulan	ulan	fulan
130. star	oha	toin	ilyówas	futuín (G)	tokul	fotuon
131. cloud	yafta	yofan (N) (T)	lak tain	añin táí (G)	lanjit ni táí	anin rano;awan
132. mist	isa?a	siwan (N)	suhan	sawan	sawan	fut
133. rain	roá	ulan	roti	uran	kadama	uren
134. thunder	dóla (A)	dola	lolas	dudun	dudun	ŋuŋur namár (E)
135. lightning	héfna (A)	kalak	yamyamat	tinitik	kanitik	ndatarák (E)
136. wind	simala	yakin	yakin	añin	añin	anin
137. blow	(k)uma	nu:	hosúf	utun	mufa?	mbu+furuko
138. hot	hasa-	bafanat	bobit	mfnas	kuan	wanato
139. cold	mori	ba?idik	sirin	rifi	lidin	rindin
140. dry	malafu	mamala	momala	kirimas	-maha	gelekár
141. wet	boro	loba loban	mareman	kiribotan	nakabetan	Jinár
142. heavy	meféla (A)	biat	tonit	maleman	kaleman	fafot;fafót
143. fire	waha	yaf ^y	yah	afi	afi	au
144. burn	tutki (A)	dun;lut	tuin	tunu	dunu	ndurin
145. smoke	foyona	numik	yah kuba?	kubur	afi kubun	foko
146. ash	kalafuá	lav	bala hua? taín	au tai	kamtái	aftai
147. black	metena	mametan	metan	metan	maketan	metemeten
148. white	put a	babut ^y	buti	futi	mfutik	noitigo
149. red	lalahana	dadau	kasahan	mera	gilala	moro-moro
150. yellow	rarakata	unin	si?init	kuniŋ	kunin	kunil (E)

	Seti	Bobot	Masiwang	Geser	Watubela	Banda
151. <i>green</i>	kokolona	iʃo	biri biri	biri (K)	uŋ-uŋ	kairanranoko
152. <i>small</i>	mattik	fofin yana	kongon	tutuin	sikit ha	doŋuro
153. <i>big</i>	boto-	ayak	isa	bobak	leha	ra
154. <i>short</i>	kes kes	wetuwetik	karokat	kubut	towól	noitfk
155. <i>long</i>	kóila	otak	wotat	malás	malá:	naru
156. <i>thin</i>	marni	minifit	manihit	banifis	manifi	munít
157. <i>thick</i>	futulu	vitól	kuta?	batolu	batolu	karát
158. <i>narrow</i>	keneʔat	kofa-kofat	koru-korup	korok	mokot	ndetéƙ
159. <i>wide</i>	ferehana	tiban	tiha?	galawa	galawal	lafár
160. <i>ill</i>	malaʔu	mayaut	kaliʔ	masinat	mahigat	nelet
161. <i>shy</i>	mukai	mayan	maliŋ	-	-	momai
162. <i>old</i>	mutua	mi tua	ronaman	matu:;ual	-mtuka	motua-
163. <i>new</i>	folla	vovoi	hosihosin	wouwou	oku-oku	feferúŋ
164. <i>good</i>	holo	kamuin ¹⁵	hia	babán;gafin	baik;fi	bobai-
165. <i>bad</i>	lal saʔa	kaʃahatan	lalan sa	garata (P)	-	ʃahat;setan
166. <i>true</i>	leke	mam	katoti	katotu	moson	mololo
167. <i>night</i>	kahana (A)	mamenin	mesat	garan	dodan	moŋjia
168. <i>day</i>	boti	matabot	roman	dodan	dodan	refok
169. <i>year</i>	naha	na:	nala?	ŋarak	ŋarak	taun
170. <i>when</i>	bothila	bot fila	rom hila	dodan fis	dodan fihí	refilo
171. <i>hide</i>	ubafun (A)	basavunin	tainunin	wunin	bunin	mbo+moroko
172. <i>ascend</i>	saʔa (A)	sa:	sa:	saka	haka	ŋʃaka;leyan
173. <i>at</i>	ni	vi	ra (?)	bua	bo	wa
174. <i>inside</i>	loi fallal	vilalan	ramuli lalan	bua lomín	bo lomi	wa rarón
175. <i>above</i>	-	lete	leta	ata	ata	tutuno
176. <i>below</i>	-	na val	laho	wawa (?) (G)	loaka	fofán
177. <i>this</i>	anín	li	mehéi	ira	ne	i:
178. <i>that</i>	amaʔun	lin	mahóvmasó	ira	ine	iko
179. <i>near</i>	hikut	wetuwetik	kekebat	ranik;rapat	lan-lani	runfk
180. <i>far</i>	kólla	otak	wotat	rau	lau	rau
181. <i>where</i>	niamba	matawaya (N) (T)	yataúa	na-ingí	kotnáu	mbe
182. <i>I</i>	yaʔa/ya	ya	yaʔaho	aku	ak	ak/-ŋ(u)
183. <i>you</i>	au/u	ana	aisó	kau	ka	ka/-m(u)
184. <i>he/she</i>	ia/ei	ia;i	iso	na;i	i	i/-n(o)
185(a) <i>we+2</i>	ita	ita	-	kita	kita	kito/-na
185(b) <i>we-2</i>	ami	am	msoa	kam	kami	kam/-mam
186. <i>you(pl.)</i>	om	em	imahoa	kumu	kemiu	kem/-mi
187. <i>they</i>	sia/si	sia	lelsoavlelso	sidi	ila	si/-ni
188. <i>what</i>	esa	safa	nahain	la	afa	ansá
189. <i>who</i>	sia	sema	isi	sei	heha	se;sain
190. <i>different</i>	sáikin	vi lul (N)	hosin	wown	haulot	fasiri
191. <i>all</i>	sefu	yalik (N)	ohi	ababis	nusik	abisi
192. <i>and</i>	tu	dua	tama	tura	hel	ndofon
193. <i>if</i>	kalu	Ø	kalu	Ø	Ø	naku
194. <i>how</i>	hoba	dawaya	haʔaí	manlá	ufanafá	ŋenbe
195. <i>no</i>	teʔula/tewa/te	sakali	eéyi	tei	teiʔ	ta
196. <i>count</i>	neke	(rekiŋ)	reʔin	reʔen (P)	reyn (K)	neke
197. <i>one</i>	esa	san	ain a	sa	ha	sa
198. <i>two</i>	nlo	lua	ain lua	roti	lua	ruo
199. <i>three</i>	ntol	tol	ain toli	tolu	tolu	telu
200. <i>four</i>	hata	fet	ain fat	fat	fata	at

NOTES TO THE APPENDIX

1. *wade in neck deep water.*
2. (A) and other dialects have tua- for 14 and tai for 16.
3. In (A) as well as Salas ata- refers to animal liver; ata was not recorded in Seti itself.
4. In Seti fufu-, which was recorded for 24, may refer to the top of the head only.
5. of cigarettes.
6. But we note hiyowo *wed* and basowan *son-in-law*.
7. Note fakhá: *to wed*.
8. Here sawa- is taken from the (A) data.
9. Reference terms only.
10. Note, however, nunu *to extinguish a lamp*.
11. Both Geser (P) and Watabela have gul *grease*.
12. In Elat [ria] seems to be a variant pronunciation.
13. *intestinal worm*.
14. *tidal pool*.
15. See no.3 *right hand*.

THE GENETIC RELATIONSHIPS OF PHILIPPINE LANGUAGES

R. David Zorc

0. BACKGROUND

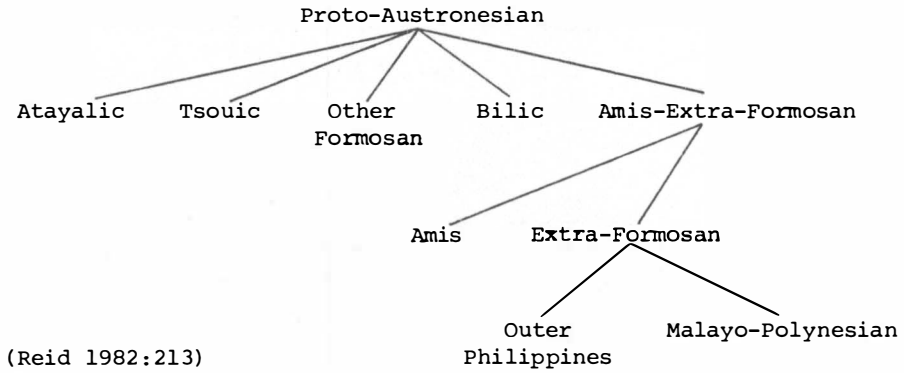
At TICAL (the Third International Conference on Austronesian Linguistics, Bali, 1982) Reid challenged the assumption that most, if not all, of the languages spoken in the Philippine archipelago descend from an immediate proto-language (PPH). The evidence he presented suggested that those languages which share in nasal infixation into root words (*CV_nCV[C]) form a subgroup of AN - the northern Philippine and Bilic languages were shown not to have *_nC cognates of PMP etyma and were thus excluded from the subgroup (Reid 1982:204,211ff). Reid therefore assigns to nasal infixation the status of a highly qualitative phonological innovation, which subgroups Central and some other Philippine languages with Malay/Indonesian and Oceanic languages. He tentatively drew Tree 1 (1982: 213) and Tree 2, which delineates the various Philippine subgroups (p.c.).¹ An analysis of the latter tree reveals a genetic chasm between Tiruray and Manobo, Kalamian and Palawan, or Bisayan and Ilokano. Tagalog is portrayed as closer to Malay than to Bontok, Cebuano to Fijian than to Kapampangan, and Sambal to Amis than to Mamanwa. A corollary to this hypothesis is that all of the exclusively shared agreements amongst Ph languages are "the effects of thousands of years of language convergence" (p.c., Reid to Ruhlen, 27.8.1982) or, those that are genuine must be relics or retentions attributable to PAN or pre-PMP. These genetic implications must be tested.

It is my purpose here to show that there are a compelling number of lexical and other innovations that substantiate a Western Austronesian node more traditionally thought of as "Proto-Philippine". Because this subgroup has a high order of diversity, and due to the propensity and natural probability for lexical replacement, not all groups continue to share all innovations. Indeed, as my lexical study has continued over the past 15 years, I have been impressed by innovations that skip over micro- or lower-level subgroup boundaries and yet delineate the same macro-subgroup established by widespread innovations. These *selective innovations* do not fit a convergence hypothesis, and hence form an integral part of my paper.

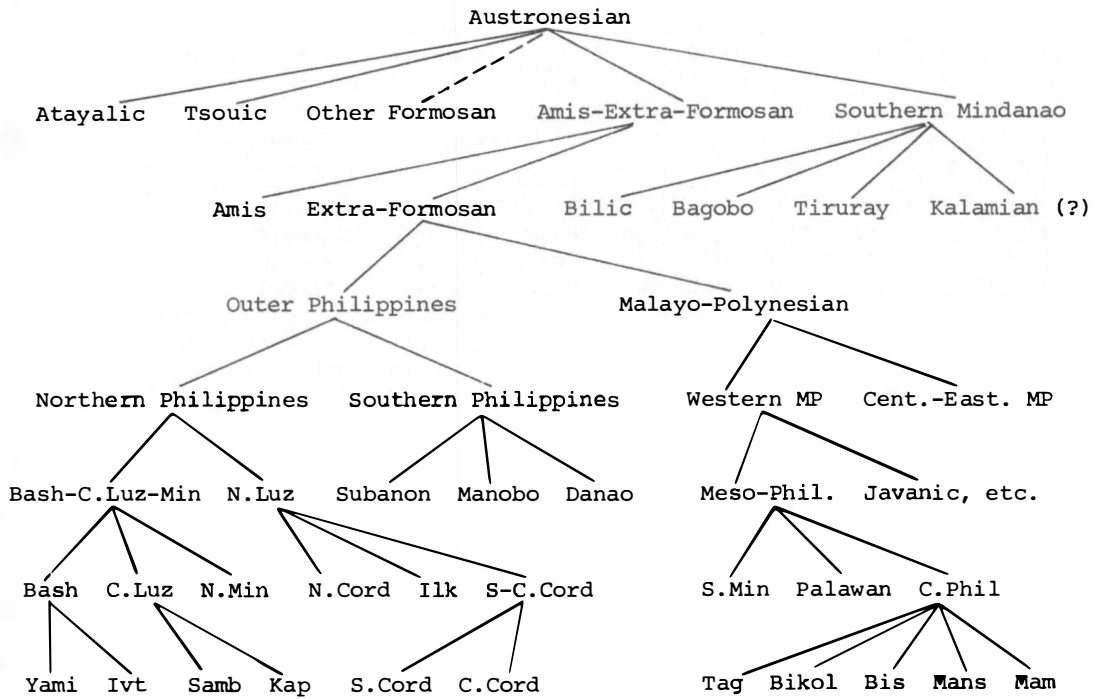
Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 147-173. *Pacific Linguistics*, C-94, 1986.

© R. David Zorc

147



Tree 1



Tree 2

Austronesian high-order subgrouping (Reid, 21.11.1981)

1. PROCEDURES FOLLOWED

(a) I began this study by organising a looseleaf edition of McFarland 1977 on the basis of the number of etyma and uniques, e.g. M5 *brain* shows only cognates of PPH *hútək *brain*, M19 only PPH *dílaq *tongue* - these were assigned a code 1-Ø (one etymon, no uniques). M36 has cognates of PPH *kukúh *finger nail*, but KnKS has a unique, kiwí - this was assigned a code 1-1 (one etymon, one unique). This was continued until my last entry was M282 *lie (falsehood)* - Ø-22 (no [widespread] etymon, 22 uniques [see 3.7 below]).

(b) I then compared these with data in Reid 1971 and Yap 1977, making a card for each reconstruction that could be assigned to "PPH" or any higher level up to PAN. Examples of the resultant cards are:

<i>eye</i> PAN *maCá > PPH *matá			Cha oho(s)
M- 07 [1-1-Ø]			Sina bulíga?
Y-253 [1-1-4]	CCrN *ʔatá		?Bot mulátot
R- 96	Mn *berən		Tas ʔighʔa
B- 45 [12345]			
D- 03 [77.2%]			
F-206 [1+3?]			
Tree: PAN > PPH > PNP/PSP			
Note: continuance into all major subgroups.			
<i>fire</i> PAN *Sapúy > PPH *hapúy			Cha fuwego
M-159 [1-Ø-Ø]	NMg *báRah		Mam gabok
Y- 20 [1-5-2]	Bs+Bk *kaláyu		Tag ^x ʔapóy<SLz
B-143 [12345]	Mk *ʔatulun		
D- 46 [15.05%]	IB1 *lipo[]		
F- 88 [1+4?]	Sn *putu		
R-112 [1+1+1]	Gor+Moŋ *súluq		
Tree: PAN > PPH > PNP/PSP			
Note: SPH intrusions with multiple replacements (except Pl, Kl, SMg, Sb, Dn, Mb, OBl, Mn)			
<i>left</i> PAN *ka+wiRí > PPH *ka+wiRí			Cha iskierda
	Iv *g(a)-uri		Gad dawí
M- 32 [1-5-4]	NCr *dimiRí		Btd wodhoŋ
Y-171 [1-5+-5]	Ilk+CCr *kǎ(N)+tigíd		Bot ʔúki
R-161	CCr *ʔigíd		Sbl ^x wiRí
B- 02 [1235]	SPh *()-ibaŋ		Sina háyin
D- 91 [8.17%]	CPh/SPh *waláh		Kpm ^x kayli?
F-383 [1+4+]	Gor+Moŋ *kələ+wiRi		Ttb ^x kaʔmbiri
			Tsw koloʔdi
Tree: PAN > PPH > PNP/PSP			
Note: multiple replacements in several major subgroups, such that there are selective retentions in Cas, Mlw, Png(NPh) and Sn, Mn(SPh).			

<i>tongue</i> PAN *Səma	
"lick" PHF *dɨlaq >	PPN *dɨlaq <i>tongue</i>
M- 19 [1-ø-ø]	PPH *dɨlaq
Y- 97 [1-ø-2]	
R-323	Cha leŋwa
B- 32 [13]	Ilt limut
D- 31 [25.8%]	
F-239 [2+5?]	Fm *li[d]am
Tree: PAN *Səma > PPH ø (loss); PHF *dɨlaq	
<i>lick</i> > PHN <i>tongue</i> (sem.shift) > PPH > PNP/PSP	
Note: continuance into all major subgroups after semantic shift, with apparent loss of PAN etymon.	

By this method, the earliest known etymologies were established, and subgroups that differed from these in form or meaning were set up. The code was expanded to show: PPH/PAN etyma - subgroup etyma - uniques and loans.

(c) I went through all the etymologies reconstructed for discrete Ph subgroups to verify and/or expand the list of proposed innovations. There is a slight imbalance in these results. Studies by Sneddon of Minahasan (1978) and Sangiric (1984) or by Reid of Central Cordilleran (1974) employ quite a large data base, while other works [e.g. by Allison of Danaw (1979), by Gallman of Mansakan (1979), or by Thiessen of Palawan (1981)] are limited to the SIL 372-meaning list. Hence, *number* of innovations should not be taken as indicative (at this stage) of historical conjectures (e.g. time-depth of separation, status as an "innovating" language, etc.) beyond the subgrouping hypotheses presented.

(d) I checked all posited innovations against all available AN reconstructions, withdrawing those that can currently be proven to be relics (selective retentions). Further research will undoubtedly reduce the various lists. However, my own research is still far from complete and more candidates will be forthcoming. Some innovations may well be shifted *upwards* (e.g. PPH < PHN or PMP), but could still be of value unless they can be established at the PAN level.

Scholars' continuing additions to AN etymologies (and Blust is to be commended in particular for his untiring efforts) increase our corpus of reconstructions to beyond double that of Dempwolff's time. Although few have been given in support of specific subgroups, the time is nigh when many of them could be. Of Blust's 443 "Austronesian etymologies II" (n.d.), only 50 were cause for withdrawal from or upwards revision of my entire list (but rarely more than three for any proposed subgroup within the entire Philippine family); a few are now put forward as PHN innovations (Table 3), along the criteria outlined below (in section 4).

2. METHODOLOGICAL PRELIMINARIES

Although there are problems involved in the isolation of lexical innovations, there are means of dealing with them. I outlined and applied certain precautionary measures (Zorc 1977:234f) and refined them (Zorc 1982:313f) as follows.

2.1 Limit forms to basic vocabulary and avoid items of trade or culture that could freely pass from one language to another

I have almost exclusively confined this study to the basic meanings found in McFarland 1977, Reid 1971, Yap 1977, Ferrell 1969 and Ray 1911, and checked these against my own data files and the data and reconstructions published by Dempwolff, Dyen, Blust, Tsuchida, Nothofer, Sneddon, Mills, et alii. [Forms considered non-basic or cultural are marked with a code "6" in the tables; see also section 4.]

2.2 Dismiss forms with phonological irregularities

That is, *not* in conformity with the standard reflexes worked out for a given language, e.g. h in a language that loses *H or *S, r in a language where *R > g, *R > y, *d/*D > d, etc. Note that *morphophonemic changes* (e.g. metathesis, syncope, assimilation, etc.) are *not* taken as phonological irregularities, and *may* be treated as innovations (e.g. Mn+Sn *leRe? *neck* < PHF *liqəR) provided that any such systematic changes cannot be counted more than once. That is, one lexical item may be included as representative of the phenomenon establishing a phonological (rather than lexical or semantic) innovation, e.g. Bs *Cl < PHN *lC; Sn *s...t < PMP *t...s, Buh -wa- < PAN *-a(qSH?Ø)u-, etc.

2.3 Reconstruct, wherever possible, an etymon for a given meaning at the earliest possible stage

For example: *blood* was PAN *Da:Ra_q, so PSP *duRú_q or SMb *dipanug *blood* are well-established innovations (but see 2.5). The 49 reconstructions in Blust 1981a (coded "12345") that have cognates in all major AN subgroups or the 51 meanings in Dyen, James and Cole 1967 that have a retention rate above 14% were considered highest in quality, and were coded "1" in the tables (viz: highly qualitative innovations). Where other etyma can be established with confidence (in both form and meaning) at the PMP or PAN level, and appear to have been replaced, they were coded "5" in the tables (see 4).

2.4 Consider the character and quality of each proposed innovation, including its geographical and linguistic distribution, potential spread, etc.

While it is difficult to distinguish a common from a spread innovation, and, in the case of conservative phonemes to isolate a borrowing, linguistic geography (such as McFarland 1977) greatly assists in showing how forms may stay within or creep over proposed subgroup boundaries.

In determining the quality of an innovation, several criteria will be discussed in section 4. Two, however, are noteworthy. Dyen, James and Cole (1967: 168) suggest that meanings of lower productivity (or low retention rates) contribute more information than those with high retention rates. The common retention of PAN *limá *five* (D1 - first on Dyen's list) or PAN *maCá (D3) indicates little more than that the languages compared are Austronesian. But, the sharing of Iv *um>tək by Ivt and Itb, or of Mk *balləŋ by Mansakan dialects in the meaning *play* (D196) or of IMb *hagsil or Iv *rukməl in the meaning *cold* (D183) *should* be highly informative and therefore indicative of subgrouping. However, I have

noted that since these meanings are highly subject to replacement, they are equally highly susceptible to borrowing. In evaluating such innovations, critical judgments become necessary.

More important are those meanings that have a high probability of retention but are nonetheless replaced. These appear to yield highly qualitative innovations, such as Mn *bərən *eye* (D3), Bk *gədaʔán or Iv *līman *die* (D7), or Bs+Bk *kaláyu, Mk *atulun, IBl *lipo[], Sn *putun *fire* (D46).

2.5 Determine if the innovation is a formal or semantic one, and if the change could happen independently

Often an old form can be established as having changed meaning (PMP *bǎRəq *to swell (as abscess)* > SBs *bagáʔ *thick*, PHF *báRaH *embers* > Iv *red*, NMg *fire*, PHN *sídəŋ *blinded by glaring light* > SMg *day*). Note that each of these semantic shifts involves the replacement of well-established etyma: PMP *(ma)kapál *thick*, PMP *ma-iRaŋ *red*, PAN *Sapúy *fire*, PAN *qal(ə)jaw *day* respectively. New forms have also been coined from previously unknown material (e.g. NCr *busáli *abscess*, PHN *lú(n)tuq *cook*), while others may be the result of reshaping (e.g. PPH *bulbul *feather* < ?PMP *bul(u)-bul(u), PHN *laqlu *pestle* > PHF *qaSəlu). When the same innovation has happened independently [Ilk nala-bága *red*, Bon ballaan *red yam* :: Iv *red* (above), or Tsw baha *fire* :: NMg *fire* (above)] further evaluation, including a re-look at the semantic assignment of the etymon, is necessary. This process could involve either rejection of the candidate(s) or a devaluation of the quality assigned.

Even after applying these measures, any proposed innovation may be a relic lost everywhere else, or as yet undiscovered in another language. However, as the number of such candidates increases, there is a strong probability that the majority will survive even protracted research for outside cognates. I take heart in the fact that of the 85 innovations posited by Reid (1974) to establish the CCr subgroup (or groups within CCr), only 18 need qualification or revision, e.g. ten may have been borrowed by neighbouring Kly, Kyp, Ibl, or Gad groups. Meanwhile, my own research has uncovered additional forms that bring the CCr total to 98; these will be published in subgroup-specific studies continuing the present paper.

3. TERMINOLOGY

3.1 Highly retentive cognates

Highly retentive cognates are found in most (if not all) major subgroups under discussion. Examples amongst Ph and AN languages include: PAN *Zálan *road, path*, PAN *qaCáy *liver*, PAN *súsu *breast*, PAN *matákut *to fear*, PAN *Cáŋis *to cry*, PAN *káʔən *eat*, PAN *ŋájan *name*, PAN *búlan *moon*, PAN *quZán *rain*, etc.

3.2 Selective retention cognates

Selective retention cognates are limited to a single subgroup (that can be established along comparative arguments) but are then found in distantly related languages. For example, WBs *dahíʔ *forehead* < PAN *daqis is limited to all

members of the proposed WBs subgroup, yet no cognate has been found in any other Ph language researched to date. Other interesting examples include:

- Dn+Mb *ʔatiŋ *perspiration* < PMP *atiŋ [Blust 1980:#24]
 Dn *nipay *snake, serpent* < PMP *nipay [Blust 1981a:#106]
 Iv *sayap *to fly* < PHN *sayap *wing* [Blust n.d.:#355]
 Bl *ʔikŋ *tail* < PMP *i(ŋ)kuŋ [Blust 1980:#167]
 PSP *qəbəl *smoke* (Mk,Sb,Dn,Mb,Bl,Pon,Tsw) < PHF *qəbəl [Tsuchida 1976]
 PNP *dánan *span [8 inches]* < PMP *dánan [Blust 1980:#442]

I have over a hundred additional examples in my files. Such etyma would not be allowed in the strictest applications of the comparative method because they are clearly retentions (not innovations). Nevertheless, (a) they serve as a synchronic isogloss around the proposed group in many observed instances, and (b) their retention amidst heavy pressure for innovation (as evidenced by replacement in the other groups) cannot be adequately explained but surely reflects some historically relevant phenomenon. I have not included these intentionally here, but numerous discoveries of similar forms will probably be forthcoming from the candidates in the various tables. However, each list in toto could still be used as a subgroup identifier for newly discovered speech varieties with some assurance of success. [See, for example, that in Zorc 1972:125-128 for WBs, or Zorc 1977:269-276 for Warayan and CBs.] This mixture of synchronic and diachronic material for subgrouping purposes is not without precedent: lexicostatistics counts the sum of retentions and shared innovations without distinguishing between them, and yet is of some value in formulating a subgrouping argument - especially if it coincides with the results of other methods.

3.3 Widespread innovation (w)

Cognates are limited to numerous language groups representing the most diverse nodes of a proposed tree, e.g. Table 1.

3.4 Selective innovation (s)

Cognates are found in only a few language groups that are geographically and genetically diverse, representing distinct nodes of a proposed tree, e.g. Table 2. Because of the distance (temporal and spatial) between the language groups involved, I can not see how a hypothesis of borrowing or convergence can be put forward apart from the proto-language immediately shared by those groups. The only other alternative is that the etyma in question are selective retentions, which may be proven by the discovery of outside cognates. I feel reasonably confident that a sufficient quantity will stand the test of time and continued research.

3.5 Contact innovation (c)

Cognates are found in two (or more) genetically distinct languages and are the result of common innovation after contact between the groups. Tag páwis and Kpm páwas clearly reflect a SLz innovation *pəwəs *sweat*, but do not serve to subgroup these languages together. Such developments help establish the degree of convergence between languages and must make the researcher chary of

positing selective innovations when *the distribution of forms* will not otherwise warrant such reconstruction, viz: a *genuine* cognate between Kpm and Tag would normally yield a PPH etymon. See Pallesen 1977 for examples of Sama-Tsg convergence.

3.6 Borrowed innovation (b)

Cognates are found in genetically distinct languages, but irregularity of form or distribution suggests in a straightforward manner that one speech variety has borrowed from another. Ntg, Agy *tambək fat* (expected ***tambəʔ*) must be loans from Bs **tambək* because of the irregular reflex for **k* (> Kl \emptyset) and the additional evidence of Kal *linuk fat* (which is probably itself a Kl innovation replacing PPH **tabáʔ*). Since virtually all Mb languages retain reflexes of PAN **Sapúy fire*, Dbw *káyu* and Ags *kaəduʔ* must be loans from a SBs dialect of Bk+Bs **kaláyu fire*. Such judgments must be made explicit, since they are open to criticism. Pallesen (1978:92f) was quite correct in his comments on my treatment of several Bs innovations:

When, however, a large number of putative exclusively shared innovations are demonstrated to have cognates outside the subgroup, then the boundaries of the subgroup are very much in question It is undoubtedly valid to identify similar forms in language B as borrowings from language A which is known to be influential, but *not* if the forms themselves are the main evidence of the influence of language B.

Exclusion of all of the forms Pallesen cites would have resulted in a more "air-tight" Bs group, but would have been tantamount to "sweeping the problems under the rug". As it is, it is probably best to list all data, facts, and hypotheses so that scholars may productively engage in debate.

3.7 Uniques (u)

Forms are limited to a single speech variety (or dialect group). Such forms do not enter into this survey, but it is noteworthy that genuine uniques are not nearly as numerous as they at first appear. Many are selective retentions. Thus, Ilt *pəsit blind* is cognate with Bin *pəsit* < PHN **pəsit* [Blust n.d.: #278], Ilt *pandək star* with Miri *fatak star* (and possibly Moŋ *pandok spot, speck* < PHN **pandək star; speck (of light)* [Blust n.d.:#260]). Ilt *tambian five* is a semantic innovation from PHN **sa+n- one + *bíləŋ count* - nor is Ilt unique in having a quinary numbering system [Dahl 1981a:fn.5, via Blust]. Sneddon (p.c.) has encouraged me to treat San and Snl as a single witness since they form a dialect chain; r or h reflexes < **R* are evidenced in dialects of each.

4. CRITERIA: TYPE, QUALITY, AND NUMBER OF INNOVATIONS

It may seem a commonplace, but it is an often overlooked fact that *every innovation means something*. Interpreting each innovation requires isolating the *types* outlined above: widespread, selective, contact, or borrowed. Furthermore, the *quality* of an innovation must be assessed. I suggest the following measures as a *rule of thumb* to weigh the innovations proposed in the various tables. The

ordering should be considered as *relative* (not absolute), so that the higher the rating, the higher the *overall* quality, e.g. 1 is better than 5, but it does not follow that 3 is *necessarily* better than 4, etc.

- 1 - Replaces a well established PAN or PMP form in a highly-retentive meaning. [See 2.4 and 3.1.]
- 2 - Formal innovation (e.g. a change in or addition of morphological material) not attested outside the group.
- 3 - Semantic innovation (i.e. the *form* may be quite old, but a definite shift in meaning has occurred replacing the etymon most closely established as having that meaning). [See 2.5.]
- 4 - A phonological or morphophonemic innovation. [See 2.2.]
- 5 - Replaces an earlier (PAN, PMP, PHN) form, but in a meaning of low retention rate or with a high probability of replacement. [See 2.4.]
- 6 - Although an item of trade or culture, distribution suggests a special kind of innovation is involved (e.g. semantic shift, pre-historical contact, etc.).
- 7 - Currently known distribution suggests innovational status, but continued research is required to establish this, viz: "none of the above".

Reactions from several colleagues have made it clear that these seven categories are perhaps better characterised as *kinds of innovations* to which further judgments concerning *quality* must be added, e.g. H (high) ... L (low). I agree that a complex morphophonemic innovation such as cluster metathesis (4H) is of greater significance than a mildly deviant semantic shift (3L) or the addition of common affixes such as *si or *i [name markers] to pronominal stems (2L). Even if this requires "going back to the drawing board", scholars should attempt to rate the quality of proposed innovations and make their own criteria explicit.

Quantity, while relative to the state of current research, must support any subgrouping hypothesis. It would be remarkable indeed if a genuine subgroup left its evidence in one linguistic area (e.g. phonology) but not in any other (e.g. lexicon or grammar). The subgrouping proposed by Reid (1982) based on nasal infixation (or the lack thereof) suffers by its singularity and the lack of additional supporting evidence. The appearance of diverse cognates of *ŋC forms could be a selective retention, a contact innovation (Ml/In influence in the central and southern Ph has been strong - see Wolff 1976 for numerous examples), or a borrowed innovation. Conversely, the non-appearance of cognate *ŋC forms could be the result of independent loss, or complex parallel developments (in Formosa and in the northern Ph or Bilic). In any event, failure to share in an innovation is not of itself proof of exclusion from a subgroup.

The origin and status of *ŋC forms needs further study and evaluation. Of the 22 forms cited from Bontok which do not show a medial nasal cluster (Reid 1982:205f), only four have CPh and SPh cognates which unequivocally reflect a nasal (*ampil *favour one person over another*, *dampilas *cliff*, *kəmpit *press, clamp*, *kindat *wink, open up eyes*) whereas eight have not been observed in these latter groups with any nasal (*apu *grandparent/child*, *ma-hátaq *unripe*, *hútek *brain*, *lútuq *cook*, *tábun *cover up*, *tahép *winnow*, *t'ipun *assemble*, *túbuq *grow*). Thus, CPh and SPh languages are intermediate on a cline between heavy nasal infixation (Oc/Ml) and little to nil nasal infixation (NPh/Fm). Note that

NPh languages do have nasal clusters in etyma that are not likely to be loans: PNP *andu *long* (by syncope < *anaduq), PNP *hiŋpis *thin*, Iv+NCr *tuŋduq *point*, ICS *suŋbat *answer*, NCr *siŋpət *good, kind*, Ilt pandək *star* (see 3.7). Nasal infixation (or its loss) is far from being established as a highly significant qualitative innovation.

5. THE EVIDENCE FOR A PHILIPPINE SUBGROUP

The number of exclusively shared lexical innovations that I have gathered thus far suggests that the languages of the Philippine archipelago (exclusive of the Sama-Bajaw group) form a single AN subgroup. This "Philippine" (or a less geo-politically prejudicial label of *Eastern Hesperonesian*) Group includes Yami (of Botel Tobago Island, within the Bashiic/Ivatanic subgroup of NPh) and the languages of northern Celebes (including Minahasan, Sangiric, Mongondow, and Gorontalic within SPh).

Widespread innovations in support of this group are presented in Table 1, and selective innovations in Table 2. Constraints of both time and space have not permitted the inclusion of the data (which can be found in the sources cited), but languages or subgroups that have cognates of the etyma under consideration are listed. The format adopted gives the following information:

- etymon number
- type of innovation (widespread, selective, etc.)
- quality [kind] of innovation, using a numerical code (section 4)
- level of reconstruction (e.g. PPH, PHN)
- reconstructed shape
- semantic assignment
- data sources (McFarland, Reid, Yap, Ferrell, Blust, etc.)
- subgroups or languages that have cognates
- (irregularities of any kind) [e.g. (+Is_g) = form probably borrowed by Isneg.]
- languages reflecting semantic shifts are put after a semicolon, along with the meaning [e.g. ;Akl *stop* = the cognate means *stop* in Akl]
- [any additional information]

I am reluctant to draw a tree at this stage and feel that Ph developments were more like amoebic colonisations than absolute splits. In general, I subscribe to the tree drawn by McFarland (1980:62) for the *upper nodes* (viz: PHN > PPH > PNP/PSP) and to that drawn by Reid (included herein as Tree 2) for the *lower nodes* (e.g. NPh includes [Iv+SLz+NMg] + [NCr+ICS]; SPh would include Bl+[Mb+Dn+Sb][CPh+SMg+Pl+Kl], Moŋ, Gor, Sn+Mn).

Because Reid's hypothesis puts Bilic and the NPh languages closer to Formosan than to Malayo-Polynesian languages, I include Table 3 to show the affinity of *all* Ph languages to Western Austronesian (PHN), i.e. the subgroup of next highest order.

6. FUTURE DIRECTIONS

This paper represents the first edition of a study intended to bring together evidence (published and unpublished) for each Philippine micro-subgroup. The macro-subgroup (PPH or Proto-Eastern Hesperonesian) is dealt with here.

Future publications will deal with Proto-Northern Philippine and Proto-Southern Philippine, and each lower-order proto-language. An Index to all of the proposed tables is included herewith. Scholars who wish to receive a copy of lists or tables prior to their eventual publication should contact me concerning those they require.

I appreciate that most of the points and arguments discussed here will need careful study and analysis. To scholars embarking on subgroup specific studies, I would recommend the consideration of what I have said about evaluating innovations, and the methods I have adopted here concerning *type*, *quality*, and *number*. I am keen to hear from colleagues about their problems and experiences in this area, and to receive critiques (whether positive or negative) about my suggestions and methodology.

I wish to acknowledge with the deepest gratitude the many positive and helpful suggestions received from Paul Black, Bob Blust, Prof. Dyen, Mat Charles, Andy Pawley, Laurie Reid, and Jim Sneddon, which have been incorporated throughout this study.

NOTE

1. To examine the geographic location of the subgroups, the reader is referred to the two major atlases, McFarland 1980, and Wurm and Hattori 1983, maps 31-34.

INDEX TO TABLES RESULTING FROM THIS STUDY

- | | |
|---|--|
| 1. PROTO-PHILIPPINE (WIDESPREAD) | 8. THE NORTH CORDILLERAN SUBGROUP |
| 2. PROTO-PHILIPPINE (SELECTIVE) | 9. THE CENTRAL CORDILLERAN SUBGROUP |
| 3. PROTO-HESPERONESIAN (THE Ph ::
WESTERN AUSTRONESIAN CONNECTION) | 10. THE SOUTH CORDILLERAN SUBGROUP |
| 4. PROTO-NORTHERN PHILIPPINE | 11. THE SOUTHERN LUZON SUBGROUP |
| 5. PROTO-SOUTHERN PHILIPPINE | 11A. EVIDENCE FOR A SUBGROUP COMPRISING
BASHIIC, SAMBALIC, KAPAMPANGAN,
SINA'UNA, AND NORTH MANGYAN |
| 5A. BILIC :: SPh CONNECTION | 11B. THE SOUTHERN LUZON :: NORTHERN
MINDORO CONNECTION |
| 5B. CPh :: SPh CONNECTION | 12. THE SAMBALIC SUBGROUP |
| 5C. PALAWANIC :: SPh CONNECTION | 13. THE NORTH MANGYAN SUBGROUP |
| 5D. SUBANON :: SPh CONNECTION | 14. THE INATI SUBGROUP OF PANAY (in
conjunction with D. Pennoyer)
Note: PAN *R > Ati d (not g,y,r,l) |
| 5E. SANGIRIC :: SPh CONNECTION | 15. THE SOUTH MANGYAN SUBGROUP |
| 5F. MINAHASAN :: SPh CONNECTION | 16. THE CENTRAL PHILIPPINE SUBGROUP |
| 5G. GORONTALO and/or MONGONDOW ::
SPh CONNECTION | 16A. THE BIKOL SUBGROUP |
| 6. THE GREATER ILOKAN CONNECTION | 16B. THE BISAYAN SUBGROUP |
| 6A. ILOKANO :: CENTRAL AND SOUTHERN
CORDILLERAN SUBGROUP | 16C. THE MANSAKAN SUBGROUP |
| 6B. ILOKANO :: CENTRAL CORDILLERAN
SUBGROUP | 17. THE KALAMIANIC SUBGROUP |
| 6C. CENTRAL AND SOUTHERN CORDILLERAN
SUBGROUP | 18. THE PALAWANIC SUBGROUP |
| 7. THE BASHIIC (IVATANIC) SUBGROUP | 18A. THE SOUTH PALAWAN SUBGROUP |

(Index cont'd)

- | | |
|------------------------------------|--|
| 19. THE SUBANON SUBGROUP | 23. THE BILIC SUBGROUP |
| 20. THE DANA O SUBGROUP | 24. THE SANGIRIC SUBGROUP |
| 21. THE MANOBO SUBGROUP | 24A. THE NORTH SANGIRIC SUBGROUP |
| 21A. THE NORTH MANOBO SUBGROUP | 25. THE MINAHASAN SUBGROUP |
| 21B. THE INLAND MANOBO SUBGROUP | 26. THE MINAHASAN AND SANGIRIC
SUBGROUP |
| 21C. THE SOUTH MANOBO SUBGROUP | |
| 22. THE DANA O AND MANOBO SUBGROUP | |

COMPREHENSIVE LIST OF ABBREVIATIONS USED

Note: Abbreviations of language names are followed by the most common name for that language and then by an indication of the subgroup within which that language falls (in parentheses).

- | | | | |
|------|--------------------------|------|--|
| Abk | Abaknon (Sml) | Bol | Bolinaw (Sbl/SLz) |
| Abr | Aborlan-Tagbanwa (NPl) | Bon | Bontok (CCr) |
| Ach | Acehnese | Bot | Botolan (Sbl/SLz) |
| Ags | Agusan (EMb/IMb) | Br | found in Borneo |
| Agta | Agta (NCR) | Bs | Bisayan group (CPh/SPH) |
| Agy | Agutaynen (Kl) | Btd | Batad (Ifg/CCr) |
| Akl | Aklanon (WBs) | Btk | Batak (NPl) |
| Alc | Alcantaranon (WBs) | Btq | Batangan (NMg) |
| Alq | Alangan (NMg) | Bty | Bantayan (Ban/Bs) |
| Amd | Amduntug-Kallahan (SCr) | Bug | Buginese (SSw) |
| Amg | Amganad (Ifg/CCr) | Buh | Buhid=Buid (SMg) |
| Ami | Ami(s) (Fm) | Buhi | Buhi (IBk) |
| AN | Austronesian | Bun | Bunun-Isbukun (Fm) |
| Apy | Apayao (NCR) | But | Butuanon (SBs) |
| Ars | Arosi (Oc) | Byn | Bayninan (Ifg/CCr) |
| Ata | Ata (CMB/IMb) | Cam | Camotes (CBs) |
| Ati | Inati of Panay | Cap | Capiznon (CBs) |
| Atta | Atta (NCR) | Car | Caraga (Mk) |
| Aty | Atayal (Fm) | Cas | Casiguran-Dumagat (NCR?) |
| B | Blust's publications | CBk | Coastal Bikol subgroup (Bk/CPh) |
| Baj | Bajaw (Sml) | CBs | Central Bisayan subgroup (Bs/CPh) |
| Ban | Bantu'anon (Bs) | CCr | Central Cordilleran subgroup (NPh) |
| Bgb | Bagobo (SMb) [not Gia] | Ceb | Cebuano (Bs) |
| Bik | Bikol (usually Naga) | Ch | Chinese |
| Bin | Bintulu (NSr) | ChaC | Cavite-Chabacano [Sp creole] |
| Bj | Bandjarese-Malay | ChaZ | Zamboanga-Chabacano [Sp creole] |
| Bk | Bikol subgroup (CPh/SPH) | Chm | Chamorro |
| Bkd | Binukid (NMB) | CLz | Central Luzon feature |
| Bl | Bilic subgroup (SPH) | CMM | Central Mindanao Manobo =
Kiryenteken (WMb/IMb) |
| B1 | Balinese | CPh | Central Philippine subgroup (SPH) |
| Blit | Blit (SMb) | Cr | Cordilleran (NPh) |
| Blk | Bulalakawon (WBs) | CSC | CCr and SCr subgroup (ICS/NPh) |
| Blw | Balangaw (CCr) | D | Dyen publication |
| Bnt | Bantik (SSn) | D* | Dempwolff reconstruction |
| Bng | Banggi | Dav | Davaweño (Mk) |
| Boh | Boholano (Ceb/Bs) | | |

Abbreviations (cont'd)

Db	Doublet	Iwk	I'wak (SCR)
Dbw	Dibabawon (EMb/IMb)	Jaun	Jaun-Jaun (SBs)
Dj	Disjunct	J-M	Jama-Mapun (Sml)
Dn	Danao subgroup (SPH)	Jv	Javanese (modern)
Dtg	Datagnon = Ratagnon (WBS)	Kag	Kagayanen (NMB)
EMb	Eastern Manobo (IMb)	Kal	Kalamian (Kl)
F	Ferrell publication	Kam	Kamayo (Mk)
Fj	Fijian (Oc)	Kan	Kantilan (Sur/SBs)
Fm	Formosan	Kan	Kanakanabu (Fm)
Fu	Futuna (Oc)	Kar	Karaw (SCR)
Gad	Gadang (NCR)	Kaw	Kawayan (Hil/CCr)
GiaB	Baguio-Giangan (Bl)	Kawi	Kawi (QJv literary lg)
GiaS	Sirih-Giangan (Bl)	Kay	Kayan
GiaT	Tagakpan-Giangan (Bl)	KB	Karo-Batak
Gim	Gimeras (Kin/WBS)	Kbs	Kabasagan (Mk)
Gor	Gorontalo (SPH)	K-C	Kalamansig Cotabato Manobo (SMb)
Gub	Gubat (CBs)	Kdz	Kadazan
Guh	Guhang (Ifg/CCr)	Kel	Bario Kelabit (NSr)
Ham	Hamtikon (Kin/WBS)	Kia	Kiangan (Ifg/CCr)
Han	Hanunoo (SMg)	Kin	Kinaray'a (WBS)
Har	Hanglulu Kallahan (SCR)	Kl	Kalamianic subgroup (SPH)
Hig	Higaonon (NMB)	Kla	Kalinga (CCr)
Hil	Hiligaynon = Ilonggo (CBs)	Klg	Kalagan (Mk)
Hin	Hinaray'a (Kin/WBS)	Kl η	Kalanguya-Kallahan (SCR)
Hok	Hokkien Chinese	Kls	Kalasan-Kallahan (SCR)
Hov	Hova = Malagasy ex Dempwolff	Kly	Keley'i'-Kallahan (SCR)
Ib	Iban (Sea Dayak)	Kmg	Kinamigin (NMB)
Iba	Iba (Sbl/SLz)	Kml	Kamaligonon (IBk)
Ibg	Ibanag (NCR)	KnkN	Kankanay-north (CCr)
IBk	Inland Bikol subgroup	KnkS	Kankanay-south (CCr)
Ibl	Inibaloi (SCR)	Kor	Koronadal (IBl/Bl)
IBl	Inner Bilic subgroup	Kpm	Kapampangan (SLz)
IC	Ik and CCr subgroup	Kuy	Kuyonon (WBS)
ICC	(Inner (=Nuclear) Central Cordilleran subgroup	Kyp	Kayapa-Kallahan (SCR)
ICS	Ik and CCr and SCR subgroup	Lan	Lanao (Dn)
Ifg	Ifugao subgroup (CCr)	Leg	Legazpi-Bikol (IBk)
Ilk	Ilokano	Lib	Libon (IBk)
Iln	Ilianen (WMB/IMb)	Liv	Livunganen (WMB/IMb)
Ilt	Ilongot (SCR)	Lok	Looknon (WBS)
IMb	Inland Manobo subgroup (Mb/SPH) (includes: EMb, CMb, WMB)	Im	Lampung (Way Lima dialect)
In	Indonesian	Lub	Lubang-Tagalog (Tg/CPh)
Ira	Iraya (NCR)	Luba	Luba (CCr)
Iri	Iriga (IBk)	M	McFarland (1977) data (M) metathesis has occurred
Irn	Iranun (Dn)	Mam	Mamanwa (Mk/CPh)
Iry	Iraya (NMg)	Man	Manabo (CCr)
Isg	Isneg (NCR)	Mar	Maranao (Dn)
Isi	Isinai (CCr)	Mas	Masbateno (CBs)
Ism	Isamal (Mk)	Mb	Manobo subgroup (SPH)
Itb	Ithayaten (Iv)	Md	Madurese
Itg	Itneg (CCr)	Mdr	Mandar (SSw)
Itw	Itawis (NCR)	Mdy	Mandaya (Mk)
Iv	Bashiic/Ivatanic subgroup (NPh)	Mex	Mexican Spanish
Ivt	Ivatan (Iv)	Mg	Mangyan (Mindoro languages)
		Mgd	Magindanao (Dn)

Abbreviations (cont'd)

Mk	Mansakan subgroup (CPh/SPh)	PMJ	Proto-Malayo-Javanic
Mkb	Minangkabau (Malay)	PMK	Proto-Mansakan
Mkr	Makassarese (SSw)	PMN	Proto-Minahasan
Ml	Malay(sian)	PMP	Proto-Malayo-Polynesian
Mlg	Malagasy (data from Dahl)	PNC	Proto-North Cordilleran
Mlw	Malaweg (NCr)	Png	Pangasinan
Mol	Molbog (SP1)	PNP	Proto-Northern Philippine
Moŋ	Mongondow (SPh)	PNS	Proto-North Sarawak
Mn	Minahasan subgroup	POC	Proto-Oceanic
Mr	Murik	Pon	Ponosakan (Moŋ/SPh)
M-S	Matig Salug (Cmb/IMb)	Port	Portuguese
Msk	Mansaka (Mk/CPh)	PHH	Proto-Philippine
Mun	Munngello-Kallahan (SCr)	PPN	Proto-Polynesian
Nab	Nabasnon (WBS)	PSB	Proto-Sama-Bajaw
Naga	Naga (CBk)	PSF	Proto-South Formosan
Nat	Naturalis (SBS)	PSN	Proto-Sangiric
NCC	North Central Cordilleran subgroup (CCr/ICS/NPh)	PSP	Proto-Southern Philippine
NCr	North Cordilleran subgroup	Puy	Puyuma
NgD	Ngaju-Dayak	PWI	Proto-West Indonesian
NMb	Northern Manobo subgroup	R	Reid (1971) data
NMg	North Mangyan subgroup	R*	Reid (1974) data (CCr)
NPh	Northern Philippine subgroup	RD	Rungus Dusun
NPl	North Palawan subgroup	Rej	Rejang Melanau
N-S	Northern-Samar (CBS)	R-K	Rajah Kabunsuan (EMb/IMb)
NSn	North Sangiric subgroup	Rom	Romblomanon (CBS)
NSr	North Sarawak subgroup	Rth	Ratahan = Bentenan (SSn)
Ntg	Northern Tagbanwa (K1)	Ruk	Rukai (Fm)
OBl	Outer Bilic (Tir,Tbl)	Sa	Sa'a (Oc)
Obo	Obo (WMb/IMb)	Sai	Saisiyat-Tunggho dialect (Fm)
Oc	Oceanic	SaiT	Saisiyat-Taai dialect (Fm)
Odg	Odionganon (Ban/Bs)	San	Sangirese (NSn)
OJv	Old Javanese	Sar	Sarangani (SMb)
Pai	Paiwan (Fm)	Sar	Sarangani (IB1)
Pal	Palawano (SP1)	Sas	Sasak
Pan	Pandan (Bk)	Sb	Subanon subgroup (SPh)
Pan	Pandan (Kin/WBS)	Sbl	Sambalic subgroup (SLz/NPh)
PAN	Proto-Austronesian	SBS	South Bisayan subgroup
Paz	Pazeh (Fm)	Sbt	Sibutu (Sml)
PBS	Proto-Bisayan	SCr	Southern Cordilleran subgroup
p.c.	personal communication	Sd	Sundanese
PCP	Proto-Central Philippine	Sed	Sedeq (Fm)
PDN	Proto-Danao	Sem	Semirara (WBS)
PFM	Proto-Formosan	Sgh	Singhi
Ph	Philippine	Sia	Siasi (Sml)
PHF	Proto-Hesperonesian and Formosan	Sib	Sibalenhon (Ban/Bs)
PHN	Proto-Hesperonesian = Western Austronesian	Sin	Sindangan-Subanon (Sb)
PIN	Proto-Indonesian	Sina	Sina-una (SLz)
Pl	Palawanic	Skt	Sanskrit
Pl	Palau	S-L	Samar-Leyte (= Waray) (CBS)
P-M	Palun-Mapun (Sml)	SLz	Southern Luzon subgroup (NPh)
PMB	Proto-Manobo	Sm	Samoan (Oc)
		SMb	Southern Manobo subgroup
		Smg	San Miguel (Mk)

Abbreviations (cont'd)

SMg	South Mangyan subgroup (SPh)	Tgk	Tagakaolo (Klg/Mk)
Sml	Sama/Samal subgroup	Tha	Thao (Fm)
Sn	Sangiric subgroup	Tic	Ticao (Mas/CBs)
Sn1	Sangil (NSn)	Tig	Tigwa (CMb/IMb)
Soc	Siocon-Subanon (Sb/SPh)	Tina	Tina (Sbl/SLz)
Sor	Sorsogon (CBs)	Tir	Tiruray (Bl)
Sp	Spanish	Tkd	Takituduh-Bunun (Fm)
SPh	Southern Philippine = Sulic	TM	Timugon-Murut
SP1	South Palawan subgroup	Tmb	Tombulu (Mn)
Sro	Saaroa (Fm)	To	Tongan (Oc)
SSw	South Sulawesi	Tse	Tonsea (Mn)
Sug1	Sugodnon-1 (Inati of Panay)	Tsg	Tausug (SBs)
Sug2	Sugodnon-2 (Inati of Panay)	Tso	Tsou (Fm)
Sur	Surigaonon (SBs)	Tsw	Tonsawang = Tombatu (Mn)
Tag	Tagalog (usually Manila dialect)	Ttb	Tontemboan = Tompakewa (Mn)
TagM	Marinduque Tagalog	Ubo	Ubo (Tbl/Bl)
TagS	Southern (Batangas) Tagalog	UJ	Uma Juman
TAG	[See: Ferrell 1969]	Umr	Umirey-Dumagat (Ncr?)
Tal	Talud = Talodda (NSn)	Vir	Virac (CBk)
Tas	Tasaday (SMb)	WBM	Western Bukidnon (WMB)
Tau	Taubuid (SMg)	WBS	West Bisayan subgroup
TB	Toba-Batak	War	Waray (S-L/CBs)
Tbl	Tboli = Tagabili (Bl)	WMB	Western Manobo subgroup (IMb)
Tbw	Tagabawa (SMb)	Y	Yap (1977) data
Tdn	Tondano (Minahasan)	Yak	Yakan (Sml)
Tdy	Tadyawan = Balaban (NMg)	Yami	Yami (Iv)
Tg	Tagalic; Tagalog subgroup	Yog	Yogad (Ncr)

TABLE 1: PROTO-PHILIPPINE INNOVATIONS - WIDESPREAD

01. w6 PPH *ʔabaká *hemp* (MRY) Iv, NCr, Ilk (CCr) SCr, SLz, CPh, Pl, Kl, Sb, Dn, Mb (Tir)
02. w7 PPH *qínit *heat (of sun)* (MRY) Ilk, CCr, Han *sun*; CPh, Pl, Kl, Sb, Mb, Bl; Mk *sweat*, Amg *boil*, Isg *reheat*
03. w2 PPH *baŋaʔ *earthenware vessel* (MRY) [PHF *b<a>ŋaʔ] NPh *cooking pot* Iv, NCr, Ilk, CCr, SCr; SPh *water jar* CPh, Sb+SLz
04. w7 PPH *baybay *shore* (BMRYZ) CPh *sand*, SPh *shore*, NPh *sea* [Blust 1970:#36 reconstructs *baSay *bank, shore*, including Kayan bahei, Kenyah bai, but has since abandoned this etymology (p.c.); the Br forms derive from a monosyllabic stem (*bay/*b<aR>ay) and indicate a possible PHN or pre-PPH etymon on which the Ph doubled monosyllable is based]
05. w1 PPH *bulbul *feather; post-pubescent hair* (BMRSY) Iv, NCr, SLz, NMg, SMg, CPh, Pl, Kl, Sb, Dn, Mb, Moŋ, Mn+Sml; +Br
06. w5 PPH *dakól *many; big* (BMRSY) *big* NCr, Ilk, CCr, Tg, Buh, Mb, Tir, Sn, Mn; *many* CCr, SCr, Kpm, Bk, Pl, Sb, Dn, Mb
07. w4 PPH *dayúq *far* (MRY) NCr, Ilk, SLz, CPh, Sb, Moŋ, Bl [Note variety of shapes ultimately derivable from a PMP *diauq, e.g. M1 jauh, CLz *ha-dawíq, and alternate prefixes, *ha- [measure] vs *ma- [adj]]

Table 1 (cont'd)

08. b5 PPH *Rúyud *pull/drag-along* (MRY) Ilk(+CCr,SCr,SLz),CPh,Pl,Han,Sb,Dn,Mb, Moŋ,Sn+Sml
09. w7 PPH *hílut *massage* (MRY) NCr,Ilk,CCr,SCr,SLz,CPh,Han,Pl,Kl,Dn,Mb
10. w5 PPH *híwa? *cut, slice* (MRSY) NCr,Ilk,SLz,CPh,(Moŋ)Sn,Mn
11. w6 PPH *lanka? *jackfruit* (MZ) [PHN *nanka?] Ilk,CCr,SCr,SLz,CPh,Mar,Gor
12. w7 PPH *lújan *ride; load* (MSZ) NCr,Ilk,CCr,SCr,Sbl,CPh,Dn,Mb,B1,Sn
13. w5 PPH *pásu? *hot; burn(ed)* (BMRSY) (Iv)NCr(Ilk)CPh,Moŋ,Gor,Mn
14. w5 PPH *pəRsah *boil/abscess* (MRY) [PHN *piRsah] CCr,SCr,Bs,Mk,Kl,Sb,Dn
15. w7 PPH *púnas *to wipe* (BMRY) Iv,NCr,Ilk,CCr,SCr,Sbl,Bk,Tg,Sb,Dn,Mb; Chm
16. w3 PPH *sáliw *buy/sell* (MRY) Iv,NCr(Ilk),SCr,SLz,Buh,Sb,Moŋ,Gor
17. w5 PPH *i+sədá? *fish* (BMRSY) NCr,NMg,Han,CPh,Pl,Sb,Dn,NMb,Moŋ,Mn; *food eaten with rice* NCr,Ilk,CCr(SCr),Bs,Moŋ; +Borneo
18. w4 PPH *siám *nine* (MRYZ) Iv,NCr,Ilk,CCr,SCr,SLz,NMg,SMg,CPh,Pl,Kl,Sb,NMb,B1, Moŋ+Sml,Kdz (also several Bornean lgs) [pos PHN]
19. w7 PPH *sújud *fine-tooth comb* (RYZ) Iv,NCr,Ilk,CCr,NMg,Buh,CPh,Pl(Kl)Dn(Mb), B1
20. w2 PPH *taRa+qinép *dream* (MRY) Iv,NCr,Ilk,SLz,NMg,SMg,CPh,Pl,Kl,Sb,Dn,Mb,Moŋ
21. w7 PPH *tan?aw *look-(far)* (MYZ) NCr,Ilk(CCr *tam?aw),CPh,Dn,NMb,Moŋ
22. w7 PPH *táwaR *call* (MSYZ) (Iv)NCr(Ilk)SCr,CPh(Abr,Kl),Sb,Dn,Mb,B1,Mn+Sml
23. w5 PPH *tulúd *push* (BMRY) CCr,SCr,Sbl,CPh(Btk)Kl,Dn,Mb,Tbl,Moŋ,Tsw,San+Sml

TABLE 2: PROTO-PHILIPPINE INNOVATIONS - SELECTIVE

24. s2 PPH *qa?juŋ *nose* (MYZ) Ilk,SLz,Kl (SP1 *əduŋ)
25. s4 PPH *[h]a-ndu *long* (MRSY) CCr, Sbl; Sn *nandu [PAN *a-naduq (B); Kayan aro? may invalidate, but the Ph cognates show syncope, the Kayan loss of the entire first syllable]
26. s3 PPH *a+núh *what?* (MRY) Agta,Luba,Man,Itg,CPh (Moŋ,Gor *ə+nu) [PHN *anúh *whatchamacallit*; Reid treats the NPh forms as loans (p.c.)]
27. s4 PPH *?anúk *chicken* (MRY) NCr,B1 [PMP *manúk *bird*; this could be the result of independent/parallel development (Reid, p.c.)]
28. s5 PPH *?áŋas *face; forehead* (MRY) Knk,Ifg,Kly,SMg,Sb,B1
29. s7 PPH *?atúbaŋ *front* (MZ) NCr,Bk,Bs,Han,Mb
30. s7 PPH *?əlák *sound of snoring/choking* (MRYZ) Bik,Png *laugh*; Kl,Iry,Sbl, Ifg *sleep*, AKl *choke*
31. s7 PPH *?idáu [*snake*] (MRY) NCr,Han,Iry,Kl [Reid (p.c.) suggests forms meaning *omen bird* (e.g. Bon ?ídəw may also be related)]
32. s5 PPH *?íRit *nit* (MRY) CCr (also: *kílit), SarMb *delouse*

Table 2 (cont'd)

33. s7 PPH *ʔin+də[gR] *stand* (MRY) NCr *wait*; Mk, Sb [Blust (p.c.) relates to Ml *injak step, tread* < *inzeg, but other Ph evidence (Ilk takdər) suggests a monosyllabic stem *dəR, also noted by Reid (p.c.)]
34. s1 PPH *ʔípus *tail* (BMRSYZ) Iv, NCr, Ilk, CCr, SP1, Moŋ, Mn
35. s1 PPH *ʔituʔ *dog* (MRY) Atta, Sin, Tir; WMB *puppy* (Ivt *cituʔ*, Ibg *kítuʔ*)
36. s7 PPH *ʔudu[] *medicine, charm* (MRSY) NCr, Ilt, Sn
37. s1 PPH *ʔŋaʔ *child* (MRYZ) CCr (Kly, Kyp) NMg, Buh, Akl, Rom (Bl *ʔəŋaʔ)
38. s2 PPH *ʔu+piá *good* (RY) Ilt, S-L(M), Ntg, WMB
39. s7 PPH *-úsiŋ *charcoal* (MRY) CCr, Bl
40. s7 PPH *bak+bak *frog* (ERMY) Cas, CCr, Dn, Mb [possibly onomatopoeitic]
41. s7 PPH *básul *blame* (MZ) Ilk (+Isq), CCr, CPh, Mb
42. s7 PPH *buqá1 *leg-joint* (ERYZ) Bl, Mb *knee*; WBS, Bk *heel*, Tag, Ceb *ankle*, Cas *back part of knee* [Dj: Bot *boʔá* < Sbl *buʔəd (M#60) *heel*]
43. s2 PPH *bu-báhi *woman* (GRY) Kla, Mk [Note: Blw *bubáʔe* shows a regular development of *a>u/b-b, see Blw *bubʔá tooth* < *baqbaq (Reid, p.c.)] [G = Gallman]
44. s1 PPH *daʔgun *year* (MRYZ) NCr, NMg, Han, WBS, Dn
45. s7 PPH *daqtaR *floor* (MRY) Iv, NCr, CCr, SCr, Pl, Kl
46. s4 PPH *daRəm *needle* (MRY) Iv, Sbl, Mk (EMb), WBS [PAN *ZáRum]
47. s6 PPH *daŋanan *pillow* (MRY) Png, Bl (K-C)
48. s7 PPH *dáyaw *praise/honour* (LSZ) NCr, Ilk (Png), Bs, Sn, Mn
49. s5 PPH *Riduʔ *earthquake* (RSY) (Bon *gidó*), PMN *ehdoʔ [The development of *R > Bon g is irregular, and indicative of a loan, but no source language can be determined]
50. s6 PPH *Rutay *hemp* (RY) Cas, IBl, Sn
51. s7 PPH *hábuŋ *shelter* (MRYZ) Iv, Ilk, CCr, SCr, Tag, Mar
52. s5 PPH *həŋít *laugh* (RYZ) Itb, Ilt, WBM
53. s7 PPH *húRay *stop; wait* (MRYZ) Cas, Ilk, CCr *wait*; Akl *stop* [Pai *paŋuay postpone, procrastinate* < Dj *pa-Suay *spend-time*, cf. Bs *pahúway relax*]
54. s5 PPH *ka-RabíʔiH *yesterday; last night* (BMRSY) NCr (Kla), Tg, Mk, Abr, Dn, Sn [Both *ka- [past time] and *RabíʔiH *night* trace to PAN, but this particular combination in the meanings cited has thus far only been found in the Philippines]
55. s7 PPH *kalasan *forest* (MRY) Ifg, SCr, Bkd
56. s2 PPH *ka+yu *you(pl.)* (MRY) NCr, Ilk, CCr (Ibl, Png) Kpm (+Tag), Mb *kiyu (A) [The combination of both elements is unique to the Philippines]
57. s7 PPH *kəRáŋ *scab* (SZ) Ilk, SBs, Mar, Tir, Gor, Mn; Db: SPh *kəRán WBS, Han, Moŋ [Blust n.d.:#159 PHN *kuRan, but Han is cognate with this form; Iban *kurai hard, rough patches of skin* (Scott), *mottled, of skin* (Richards) < *kúray]
58. s5 PPH *kələp *night; dark* (PRYZ) Cas, S-L, Ati *night*, Mlw, Itw *dark* [cf. Iv *aʔləp *night*]

Table 2 (cont'd)

59. s7 PPH *kúRun [*cogon-grass*] *Imperata cylindrica* (BSZ) (Bon,Sgd gólon), Tag, Btk,Han,Mn
60. s5 PPH *kunəm *cloud* (MRY) NCr,Pl(Thiessen),Kl
61. s5 PPH *ləbág *to swell; abscess* (MRY) Kia,Png,Kpm(Tag),Mar,Mb
62. s7 PPH *lipéd *to hide* (RY) Isg,Kuy
63. s5 PPH *lúsi? *penis* (MRYZ) Ilt,Bon,Knk lúsi, GiaS luhi? [Note also: Han pu-sli?, Bol luspí?]
64. s2 PPH *ma- *one unit* (e.g. 10, 100, 1000) (MRY) NCr,Sbl,Mk,Sn [Possibly a parallel development (Reid, p.c.)]
65. s2 PPH *n-atáy *dead* (MRY) NCr,Ilk,CCr(Kyp,Kly)Bot(Tag)Sn [Possibly the reduction of PAN *(m<i)n>áCéy, cf. Puy mianaTag, Ivt,Cas,Kpm,Mar]
66. s7 PPH *nisnis *wipe; brush* (RSY) Iv,NCr,Mn
67. s7 PPH *ŋa?ŋa? *betel chew* (MRSY) Bot,Png,Alŋ,Iri,Tag,Mn [If SLz < Tag, then may shift to SPh innovation]
68. s5 PPH *ŋítit *black* (MRY) NCr(>Ilk,Kla,Bol),CCr,Sbl,Kuy [Db: PPH *ŋit()ŋit *dark* Knk,CPh]
69. s5 PPH *pag?uŋ [*turtle*] (MRY) Ilk,CCr,Sbl,Tag,Mam
70. s7 PPH *pantaR *sand* (RY) Knk,Kly,Iln [Db: PSP *pantad]
71. s5 PPH *pawíkan [*turtle*] (BMRSYZ) Cas,Ilk,Png,CPh,Han,Pl,Kl,Mar,Tir,Sn,Mn
72. s5 PPH *pítək *mud* (MRY) Ilk,CCr,SCr,B1 [Db/Dj: PMP *pítak (Blust n.d.:#291)]
73. s7 PPH *pú?əj *thigh; knee* (MRY) CCr,SCr *knee*, Kpm,EMb,Klg,IB1 *thigh*, Mk *buttocks*
74. s7 PPH *pulaw *hunt (at night)* (SYZ) Ivt,Bkd *hunt*, Akl,Han *stay up late*, Mn *wake, get up*
75. s7 PPH *putut *short; cut off* (MSZ) (Ivt *fracture*); Ilk,Man,Bon,Itg,Yog *offspring*, Kia,Bon,Png,Kpm,Sn *cut-off*, CPh,Mn *short (person)*
76. s7 PPH *sa?(ə)gəb *fetch water* (MYZ) NCr(CCr)CPh,Dn (Sb *sigəb)
77. s7 PPH *sak(ə)du *fetch water* (MYZ) NCr,Ilk,CCr,Kpm,Mb
78. s4 PPH *sa-siám *nine* (MRY) Ivt,Ata [*C₁a- reduplication is probably PAN, but the shape *siám is here considered an innovation, see #18]
79. s5 PPH *sə?ít *thorn* (MRY) NCr,Ilk,Itg,Man,Luba,Tg,Bs
80. s5 PPH *səkí *foot, leg* (MRYZ) NCr,CCr,Bs,Tir
81. s7 PPH *səjəb *burn* (MRSY) Gad,CCr (Sn *soRob; Moŋ turub)
82. s7 PPH *səlár *big* (MRSY) Sbl(Ifg,Blw),SBs,Soc(Dn)Mn
83. s2 PPH *si?ák I (MRY) Ilk,Png,Ilt,Mb [Ilk si?akó-n I ... *already* suggests these derive from *si-ak(ú) (Reid, p.c.)]
84. s2 PPH *sa(?)kən I (MRY) Ibg,Cas,CCr,Mar
85. s2 PPH *si-kamí *we(excl.)* (MRY) NCr,Png,Ilt,Tina,Mb [Possible parallel development]

Table 2 (cont'd)

86. s7 PPH *sukay *comb; delouse* Ilk search, Bon, Sgd put decoration in hair, SBs, Mk, Pl *delouse*
87. s5 PPH *taRənək *mosquito* (RSY) Ilt, Mk, Dn, Mb, Sn [If Ilt tənək not cognate (e.g. < *Cənək *pierce*), then shift to SPh]
88. s7 PPH *taRád *to wait* (ERY) Png, SBs, Mk(Soc)Mb, Tir
89. s5 PPH *tanud *thread; needle* (MRYZ) Kla, Isi, Yog *needle*, Ceb, Tsg, Dn(>Mb, KorBl) *thread*
90. s5 PPH *təm+təm *to burn* (RYZ) Ilk, Mb; Bl *təm
91. s1 PPH *tiʔris *urine* (MYZ) Png, Bk, Tdy; Han *stinging secretion of a millipede*
92. s7 PPH *tiŋén *look-for/hunt* (MRY) NCr, Tag, Mb
93. s5 PPH *tuqlid *straight* (MRY) SLz, Han, CPh, Kl, Mgd, Mb, Snl, Moŋ; +Br?
94. s7 PPH *tubáR *answer* (MRY) Isg, Bot, SBs, Mk(NMb), Abr, Kl [Isg tubág pos < NPh *t<u(m>ə)báR, Bot db tóbáy; pos only SPh]
95. s5 PPH *tudul *give* (MPRYZ) Iv, Kl, Ati
96. s2 PPH *tulduq *to point* (MRSY) NCr(Ilk)Kpm, CPh, Kl, Mn [Shift to "index-finger" common]
97. s7 PPH *wak+wak *crow* Cas(Png), Bkd, Mam [cf. PHN *uak *crow*; Bs *wakwak *witch*; probably onomatopoeic]
98. s? PPH *dik()ləm *night* (RY) Agta hikləm, Tbw dikilum

TABLE 3: PROTO-HESPERONESIAN INNOVATIONS - THE PHILIPPINE :: WESTERN AUSTRONESIAN CONNECTION

01. s5 PHN *qaʔduŋ *sit* (BRTYZ) Pl, Kl; Mny maharuŋ [Blust 1973:#242]
02. s6 PHN *a(m)bək *mat* (BMRY) NCr, CCr, SCr; Beta ambok [Blust 1980:#1]
03. s7 PHN *qajəŋ *charcoal* (DSYZ) NMg, Ib, Ml, Jv, TB, KB(etc.); Mn *soot*
04. s2 PHN *a+ti [deictic: 3] (BRYZ) Pl, Kl, NMg, SMg; UJ, Busaŋ, Malagasy [Blust n.d.: #23]
05. s5 PHN *ǎdég *back* (anatomical) (BMRY) NCr, CCr, Bik; Sgh [Blust 1973:#253]
06. s5 PHN *ə+mís *sweet* (BMRSY) NCr, Mb, Sb, Bl, Sn, Tdn; Mr, UJ, Bar [Blust n.d.:#92]
07. s4 PHN *ə+súŋ *mortar* (BRY) Bl; UJ, Bukit [Blust 1980:#127]
08. s5 PHN *íkəj *cough* (BMRY) NCr, NPl; Bintulu [Blust 1973:#247]
09. s7 PHN *íflu *orphan* (BEMZ) Bs, Bk, Han, Mb; Iban iru [Blust 1970:#337][E = Elkins]
10. s4 PHN *iném *drink* (BERYZ) WBS, Mk, Kl, Dn, Mb, Tir; KB, Bal, Sas, Rej [Blust n.d.: #109]
11. s4 PHN *ípən *tooth* (DRSYZ) WBS, Tsg, Tag, Kpm, Mn; TB ipon
12. s3 PHN *quntu *tooth* (generic) (NRYZ) WBS, Mk; Snd, Jv [Nothofer 1975:38]

Table 3 (cont'd)

13. w6 PHN *qútaŋ *debt* (DMRY) [widespread, possibly Ph < In/Ml]
14. w7 PHN *bálu *widow* (DMRY) [widespread Ph; Ach, Ib, Ml, OJv, etc.]
15. w7 PHN *báluŋ *provisions (things rolled-up for journey)* (DMSZ) w Ph/In
16. s5 PHN *batúk *cough* (ADNRYZ) S-L, Iry(AIŋ), Dn, WMB; PMJ [Nothofer 1975:124]
17. b7 PHN *báyaD *pay* (NMRSYZ) [widespread Ph; but note PSn *baeR, PMJ *bayar [Nothofer 1975:143], hence pos PHN *báyaR]
18. s4 PHN *bǎkǎn (BMRYZ) [negator of nominals] NCr, CCr, Yami, NBs, Tsg, IBk, Han, K-C, Tir [neg]; NPl, Kel, Kapuas *other, different* [Blust 1980:#52]
19. s7 PHN *bǎŋǎR *deaf(ened); dumb* (BMRYZ) NCr, Bk; Mkb, Bal [Blust n.d.:#49]
20. s7 PHN *bihaR *allow (to live); alive* (CMS) NPh, Pal, Buh, Mar, Moŋ, Sn *alive*; Ml *biar permit*
21. s7 PHN *búliR *cluster of fruit (e.g. ear of grain; bunch of bananas)* (DMNZ) NPh, CPh, Mar *bunch*; PMJ *ear of grain* [Nothofer 1975:127]
22. s1 PHN *bulúŋ *leaf* (DMRY) NPh; TB, Mlg [cf. SPh *medicine*]
23. s7 PHN *búŋet *angry* (BMRY) CCr, SCr, Mar, WBM; Kay; OJv(M) [Blust 1980:#81]
24. s5 PHN *buriŋ *charcoal* (RSYZ) Tir, Mn, Sn; Sml/Tsg; NgD *burin*
25. w5 PHN *bútuq *penis* (DMRYZ) Iv, Yog, Agta, Ilk, Kyp, SLz, Bs, Bk, Kl, Sb, Bl; Ib, Ml, PSS, Mlg (etc.)
26. s7 PHN *buya? *see; look-for* (RY) Ivt, Ilk, Rth; Sas, Mkr [Mills 1981:#62]
27. s6 PHN *búyu? *betel-leaf* (BRYZ) Bs, Bk, Mk, Pl, Kl; Tabun, Balait [Blust 1973:#92]
28. s5 PHN *dǎRǎŋ *red* (BMRTY) NCr, SCr, SP1; Limbaŋ, Boloŋan [Blust n.d.:#78; note Ml *jəraŋ slow heating/toasting over an open fire* and PanBk, Vir *dugǎŋ, Msk, Klq ma-gdaŋ dry* pos < PHN *zǎRǎŋ]
29. s2 PHN *di(y)a? [deictic: 2/3] (BERY) Mb, Sin(Tir); Ib [Blust 1970:#152]
30. w5 PHN *díRus *bathe* (DMNRY) Iv, NCr, Ilk, Bs, Bk, Mb, Bl; PMJ, TB [Nothofer 1975:165]
31. w3 PHN *dílaq *tongue* (DMRSYZ) PPH; Ml(M), Ib, TB, etc. [Note: PAN *Səma *tongue*, PHF *dílaq *lick* - contrast PFM *li[d]am *tongue*]
32. s1 PHN *dúdu? *breast* (BRYZ) Mas, But, Tsg, Bik, Sb, NPl; (Sml), Wolio [Blust 1980:#108] [cf. Bl *tutu? *breast*; PMJ *zuzu? *feed*]
33. s7 PHN *gak+gak *crow* [probably onomatopoeic] (DRY) Kl; Ml, Jv, Mlg [also: Abr ?u+gak]
34. s7 PHN *Rawéd *betel-leaf* (BMRYZ) NCr, Ilk, CCr, SCr, NMg; Lepu-Pohun *awat* <awət> [Blust n.d.:#330]
35. b7 PHN *Ríbu *thousand* (DMRYZ) [widespread as ^xríbu in most Ph < Ml/In; NB: Mar ŋgibo, Sb *ŋibu, Sml ibu, Jv ewu]
36. s7 PHN *ha(m)bəl *weave (cloth)* (BMRYZ) PPH; Kdz, RD, TM *weave*, Ml (h)ambal *rug, carpet* [Blust n.d.:#400]
37. s7 PHN *ha(n)di? / db: *həndi? *no, not* (BMRYZ) NCr, CCr, Bik, Bkd; Bukit, Gondaŋ / SCr, Tag, WBS, IBk, Kl, Sb; Penudjaq, Mamben [Blust n.d.:#401]

Table 3 (cont'd)

38. s7 PHN *haRəZan *ladder, staircase* (DMS;Hendon) PPH; PMN *ahdan; NgD həjan
39. s6 PHN *hasák *dibble* (BMRSY) CCr,SCr,CPh,Tbl,Sn; Taboyan, Lawangan, Dusun-Dejah [Blust 1973:#292]
40. s5 PHN *háwak *body* (DMNRSYZ) CCr,Bk,Mn,Sn; PMJ [Note: more widespread cognates with semantic shift to *waist* or *trunk*]
41. b6 PHN *kam()diŋ *goat* (BDYZ) [Widespread cognates of *kambiŋ and *kandiŋ, see Blust 1980:#173]
42. s7 PHN *kǎwá? *spider* (BMRYZ) Gad,Yog,CCr,Ibl,Kyp,Kly,Sb,Dbw,Mdy; Ml,B1 [Blust n.d.:#134; note reduplication and other formatives]
43. w7 PHN *kíday *eyebrow* (BMRYZ) Iv,NCr,Ilk,CCr,SCr,Sbl,Kpm,CPh,Pl(Kl)Sb,Dn(Mb), Kiput kira:y [Blust 1972b:#12]
44. s7 PHN *kut+kut *scratch (with claws)* (BMSZ) Ilk,CCr,CPh,Dn,Mb; Ib,M1 [Blust 1970:#221; note semantic shifts → *dig* or *grate*]
45. s4 PHN *laqlu *pestle* (BMRYZ) Ifg,Ibl,Kly,Kyp,Bot,Pl; KB [Blust 1980:#253]
46. s7 PHN *lá(m)+pis *cut-thin* (DRYZ) Agta,Cas *thin*, CPh; TB,Mlg,Jv,M1
47. s5 PHN *layaŋ *to fly* (DRY) Han,Mb,Tir(B1);Sml; TB,Jv,M1,NgD
48. w6 PHN *lǎŋáh *Sesamum indicum* (DMSZ) NCr,Ilk,CCr,Ibl,Sbl,CPh,Han,Dn,Mb,Tir; TB,Ib,M1,Jv,NgD [Dempwolff cites Oceanic cognates in the meaning *saffron*, viz: different semantics]
49. s5 PHN *lə(m)páD *to fly* (BMRYZ) Sbl,CPh,K1; M1 [Blust 1970:#247]
50. s7 PHN *lə(n)tíq *thunder storm* (BRYZ) Bs,Mk(Tag),Han,Sb,Dn,B1; LgT,Bug [Blust n.d.:#196]
51. s5 PHN *liŋʔət *sweat* (BMRYZ) NCr,Ilk,CCr,SCr,Pl,Tir,Tbl; Mlg [Blust 1980:#283] [Dj: *riŋ(a)ʔət (Z) Alŋ,Iry riŋaʔət, Sml liŋoʔot, Jv kə/riŋət]
52. s7 PHN *lúluj *shin* (DMRSYZ) Ifg,Sb,Tbl *knee*; (Ilk,Png),Bs(Tag),Mn;NgD
53. s5 PHN *lú(n)tuq *cook* (DMRSYZ) NCr,Ilk(CCr)SCr,Sbl,Kpm,CPh,Pl,Kl,Dn,Mb,Moŋ, Mn; NgD
54. w6 PHN *naŋka? *jackfruit* (DMNZ) (Ivt)NCr,Ilk,CBs,Tsg,Tag,Bik;Ib,M1,Snd,Mad, TB [Note: Kpm yaŋka? would appear to indicate *ñāŋka?] [See: PPH *laŋka?, Table 1, #11]
55. s7 PHN *pəgər *fence; enclosure* (DS) Mn; Jv,M1,NgD,Mlg
56. s7 PHN *páhíd *to wipe* (BRYZ) Bs,Bk,Tg,Pl,Kl(NMb),Gor; Kel [Blust 1970:#290]
57. s5 PHN *pálaŋ *palm (of hand)* (DMRSY) (NCr,Ilk,CCr),Sina,Sbl,Kpm,CPh,Pl,Kl, Sb,Dn,Mb,Gor,B1,Sn,Mn,Gor; Sml,TB,Mlg
58. s7 PHN *palíʔ *cut, wound* (BRSY) Bs,Agy,Sb,Dn,Mb,Tir,Moŋ,Gor,Mn; Muka [Blust 1980:#333]
59. s7 PHN *pa+saqan *carry on shoulder* (DMSYZ) CCr,Kia(Kpm),CPh,Kl,Gor,Mn;(TB)M1 [Note: Kly,Alŋ *shoulder*]
60. s6 PHN *padək *husk (of rice)* (BRY) NPl; Kiput,Loŋ-Semado [Blust n.d.:#259]
61. s7 PHN *pəkát *stick(y)* (BMRY) CCr,SCr,Mar,WBM: M1 [Blust 1973:#232]
62. s5 PHN *pəra *dry* (DS) Mn,TB,Jv

Table 3 (cont'd)

63. s7 PHN *piRsah *abscess* (BRYZ) Bk,Mb,Bl; Ib,Dalat [Blust n.d.:#288]
64. s5 PHN *pipi *cheek* (DNRSY) Bkd,Bl,Sn,Mn; Jv,Ml,NgD,Mlg
65. s7 PHN *púdut *to pick (up)* (BCSZ) CPh,Dn,Mb,Tir,Moŋ,Mn; Miri [Blust n.d.:#302]
66. s7 PHN *sa(m)báw *soup, broth* (BMRYZ) (Cas,Ilt)Sbl,Kpm,CPh,Pl,Kl,Sb,Dn,Mb,Sn; LgA, LgSan, Ml [Blust n.d.:#338]
67. s7 PHN *sawáh [*snake: python*] (DRSYZ) CPh,Pl,Mn,Rth; Ib,Ml,Jv,TB,KB,Sml
68. s7 PHN *sáyaw *dance* (BMZ) Ilk,Png,Kia,CPh,Mb,Dn; LgLabid; Uma *leap*; PSS [Blust n.d.:#356; Mills 1975:820]
69. s7 PHN *sǝjém *ant* (BMSZ;Mills) NCr,IBk,Moŋ,Pon,Mn; PSS [Mills 1975:821; db PHN *si(n)jəm Blust n.d.:#371]
70. sl PHN *sulu *finger nail* (BCESYZ) NMg *finger*; Mb,Mn; Busaŋ, Murik,Aŋ-Batak [Blust 1973:#97]
71. w5 PHN *tabó? *fat* (DMRSTY) Iv,NCr,Ilk,CCr,SCr,Sbl,Kpm,Mk,Tg,Bk,Pl,Han,Sin, Mb,Bl,Gor,Sn,Mn; TB,Mlg
72. s7 PHN *táma? *hit the target; correct* (DMZ) Ilk *suitable*, CPh,Kl,Han; Jv,Ml, Ib,Ach [Disassociated from Arabic *tamma complete*]
73. s7 PHN *təy+təy *bridge* (DEMSZ) CCr *ladder*, NCr,Sbl,Kpm,CPh,Dn,Mb,Mn,Tir,Moŋ; Ml titian, Mlg tetezana
74. s5 PHN *təŋtəŋ *see; look-at* (BRSY) Mb,Dn,Tir,Mn; Ib [Blust 1972b:#111]
75. s4 PHN *tuqlaŋ *bone* (BCDMRY) Iv,NCr,Ilk(CCr)Ilt,IBk,Pl; Sml,Rj,Ml,Ach, etc. [cf. PAN *tuqəlaN]
76. s5 PHN *tuquh *right (side)* (DMYZ) Bk,Bs(Kl); Snd
77. s5 PHN *tuqúr *dry* (CNRYZ) Sina,Kpm(>Tag),Pl; Sml; Ib tu:r, Snd tuhur [Nothofer 1975:68 associates with PMP *tuquD *to stand*, but the Ph evidence suggests a separate etymon]
78. s4 PHN *uRsa *deer* (BMRY) NCr,Ilk,CCr,SCr,Sbl(Iv); TB [Db: *Rusa - Blust 1970: #367; see PSP *?usá]
79. s7 PHN *sayap *wing* (BFMRY) Iv *to fly*; Ib,Ml *wing* [Blust n.d.:#355]

REFERENCES

ALLISON, E. Joe

- 1979 Proto-Danaw: a comparative study of Maranaw, Magindanaw, and Iranun. *Papers in Philippine linguistics* No.10. PL, A-55:53-112.

ARANETA, F. and M.A. BERNAD

- 1960 'Bisayans' of Borneo and the 'Tagalogs' and 'Visayans' of the Philippines. *Sarawak Museum Journal* 9:542-564.

BARBIAN, Karl-Josef

- 1977 The Mangyan languages of Mindoro. M.A. dissertation, University of San Carlos, Cebu City.

BLUST, Robert A.

- 1970 Proto-Austronesian addenda. *Oceanic Linguistics* 9:104-162.
- 1972a Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages. *Working Papers in Linguistics, University of Hawaii* 4/1:1-43.
- 1972b Additions to 'Proto-Austronesian addenda' and 'Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages'. *Working Papers in Linguistics, University of Hawaii* 4/8:1-17.
- 1973 Additions to 'Proto-Austronesian addenda' and 'Proto-Oceanic addenda with cognates in non-Oceanic Austronesian languages' - II. *Working Papers in Linguistics, University of Hawaii* 5/3:33-61.
- 1974 The Proto-North Sarawak vowel deletion hypothesis. Ph.D. dissertation, University of Hawaii.
- 1977 The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report. *Working Papers in Linguistics, University of Hawaii* 9/2:1-15.
- 1980 Austronesian etymologies. *Oceanic Linguistics* 19:1-181.
- 1981a Variation in retention rate among Austronesian languages. Paper presented to the Third International Conference on Austronesian Linguistics, Denpasar, Bali.
- 1981b The reconstruction of Proto-Malayo-Javanic: an appreciation. *Bijdragen tot de Taal-, Land- en Volkenkunde* 137/4:456-469.
- 1982 Review of Sneddon *Proto-Minahasan: phonology, morphology and word-list*. *Language* 58/4:921-926.
- n.d. Austronesian etymologies II. To be published in *Oceanic Linguistics*.

BLUST, Robert A., ed.

- 1981 *Historical linguistics in Indonesia Part I*. NUSA Volume 10. Jakarta: Universitas Atma Jaya.

CHARLES, Mathew

- 1974 Problems in the reconstruction of Proto-Philippine phonology and the subgrouping of Philippine languages. *Oceanic Linguistics* 13: 457-509.

CHRÉTIEN, Douglas

- 1962 A classification of twenty-one Philippine languages. *Philippine Journal of Science* 91/4:485-506.

DAHL, Otto Chr.

- 1976 *Proto-Austronesian*. Scandinavian Institute of Asian Studies Monograph Series No.15. Curzon Press.
- 1981a Austronesian numerals. In Robert A. Blust, ed. 1981:46-58.
- 1981b *Early phonetic and phonemic changes in Austronesian*. Oslo: Universitetsforlaget.

DEMPWOLFF, Otto

- 1924-25 Die l-, r- und d-Laute in austronesischen Sprachen. *Zeitschrift für Eingeborenen-Sprachen* 15:19-50, 116-138, 223-238, 273-319.
- 1926 Ivatan als 'test-sprache' für uraustronesisches *l. *Zeitschrift für Eingeborenen-Sprachen* 16:298-302.
- 1938 *Vergleichende Lautlehre des austronesischen Wortschatzes, vol.3: Austronesisches Worterverzeichnis*. Berlin: Dietrich Reimer.

DYEN, Isidore

- 1953 *The Proto-Malayo-Polynesian laryngeals*. Baltimore: Linguistic Society of America.
- 1965a *A lexicostatistical classification of the Austronesian languages*. Indiana University Publications in Anthropology and Linguistics, Memoir 19. Baltimore: The Waverly Press.

DYEN, Isidore, A.T. JAMES and J.W.L. COLE

- 1967 Language divergence and estimated word retention rate. *Language* 43: 150-171.

DYEN, Isidore and Curtis D. McFARLAND

- 1970 Proto-Austronesian etyma constituting an Austronesian cognate finder list. Yale University, mimeographed.

ELKINS, Richard E.

- 1974 A Proto-Manobo word list. *Oceanic Linguistics* 13:601-642.
- 1984 An extended Proto-Manobo word list. In *PANAGANI Essays in Honor of Bonifacio P. Sibayan on his sixty-seventh birthday*, 218-229. Manila: Linguistic Society of the Philippines.

FERRELL, Raleigh

- 1969 *Taiwan Aboriginal groups: problems in cultural and linguistic classification*. Academia Sinica Monograph No.17. Taipei: Institute of Ethnology.
- 1979 Phonological subgrouping of Formosan languages. In Paz Buenaventura Naylor, ed. 1979:241-254.

FLEISCHMAN, Eric

- 1981 The Danao languages: Magindanaon, Iranun, Maranao, and Illanun. *Philippine Journal of Linguistics* 12/1:57-77.

GALLMAN, Andrew Franklin, II

- 1979 Proto-South-East Mindanao and its internal relationships. Papers in Philippine Linguistics No.10. PL, A-55:1-52.

HALIM, Amran, Lois CARRINGTON and S.A. WURM, eds

- 1982 *Papers from the Third International Conference on Austronesian Linguistics, vol.2*. PL, C-75.

LLAMZON, Teodoro A. and Ma. Teresita MARTIN

- 1976 A subgrouping of 100 Philippine languages. In Nguyễn Đăng Liêm, ed. *South-East Asian linguistic studies, vol.2*, 141-172. PL, C-42.

McFARLAND, Curtis D.

- 1974 The dialects of the Bikol area. Ph.D. dissertation, Yale University.
- 1977 *Northern Philippine linguistic geography*. Study of Languages and Cultures of Asia and Africa Monograph Series No.9. Tokyo.
- 1980 *A linguistic atlas of the Philippines*. Study of Languages and Cultures of Asia and Africa Monograph Series No.15. Tokyo.

MILLS, Roger F.

- 1975 Proto-South Sulawesi and Proto-Austronesian phonology. 2 vols. Ph.D. dissertation, University of Michigan.
- 1981 Additional addenda. In Robert A. Blust, ed. 1981:59-82.

NAYLOR, Paz Buenaventura, ed.

- 1979 *Austronesian Studies, Papers from the Second Eastern Conference on Austronesian Languages*. Michigan Papers on South and Southeast Asia No.15. Ann Arbor: University of Michigan.

NOTHOFER, Bernd

- 1975 *The reconstruction of Proto-Malayo-Javanic*. Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde No.73. The Hague.

PANGANIBAN, Jose Villa

- 1972 *Diksyunaryo-tesauro Pilipino-Ingles*. Quezon City: Manlapaz Publishing Co.

PALLESEN, A. Kemp

- 1977 Culture contact and language convergence. Ph.D. dissertation, University of California at Berkeley.
- 1978 Review of Zorc *The Bisayan dialects of the Philippines*. *Philippine Journal of Linguistics* 9:90-95.

PENNOYER, F. Douglas

- 1979 Buhid and Taubuid: a new subgroup in Mindoro, Philippines. In Paz B. Naylor, ed. 1979:265-171.
- n.d. Unpublished field notes on the Inati dialects of Panay (Inete, Sugodnin 1, Sugodnin 2).

PITTMAN, Richard S.

- 1953 *Notes on the dialect geography of the Philippines*. Third edition. University of North Dakota: Summer Institute of Linguistics.

RAY, Sidney H.

- 1911 The languages of Borneo. *Sarawak Museum Journal* 1/4:1-196.

REID, Lawrence A.

- 1971 *Philippine minor languages: word lists and phonologies*. Oceanic Linguistics Special Publication No.8. Honolulu: University of Hawaii Press.
- 1974 The Central Cordilleran subgroup of Philippine languages. *Oceanic Linguistics* 13:511-560.
- 1976 *Bontok-English dictionary*. PL, C-36.
- 1982 The demise of Proto-Philippines. In Halim, Carrington and Wurm, eds 1982:201-216.

RUHLEN, Merritt

in progress *A guide to the world's languages.*

SCOTT, N.C.

1956 *A dictionary of Sea Dayak.* SOAS, University of London.

SNEDDON, James N.

1970 The languages of Minahasa, North Celebes. *Oceanic Linguistics* 9/1: 11-36.

1978 *Proto-Minahasan: phonology, morphology and wordlist.* PL, B-54.

1984 *Proto-Sangiric and the Sangiric languages.* PL, B-91.

SOUTHWORTH, Franklin C.

1964 Family-tree diagrams. *Language* 40:557-565.

THARP, James A.

1974 The North Cordilleran subgroup of Philippine languages. *Working Papers in Linguistics, University of Hawaii* 6/6:53-114.

THIESSEN, Henry Arnold

1981 *Phonological reconstruction of Proto-Palawan.* Anthropological Papers No.10. Manila: National Museum.

THOMAS, David and Alan HEALEY

1962 Some Philippine language subgroupings: a lexicostatistical study. *Anthropological Linguistics* 4/9:21-33.

TSUCHIDA, Shigeru

1976 *Reconstruction of Proto-Tsouic phonology.* Study of Languages and Cultures of Asia and Africa Monograph Series No.5. Tokyo.

1983 Puyuma-English Index. *Working Papers in Linguistics* 83:10-63. University of Tokyo.

WALTON, Charles

1979 A Philippine language tree. *Anthropological Linguistics* 21/2:70-98.

WILKINSON, R.J.

1959 *A Malay-English dictionary (romanised).* 2 volumes. London: Macmillan.

WOLFF, John U.

1967 History of the dialect of the Camotes Islands, Philippines, and the spread of Cebuano Bisayan. *Oceanic Linguistics* 6/2:63-79.

1974 Proto-Austronesian *r and *d. *Oceanic Linguistics* 13:77-121.

1976 Malay borrowings in Tagalog. In O.D. Cowan and O.W. Wolters, eds *Southeast Asian history and historiography essays presented to D.G.E. Hall*, 345-367. Ithaca: Cornell University Press.

1982 Proto-Austronesian *c, *z, *g, and *t. In Halim, Carrington, and Wurm, eds, 1982:1-30.

WURM, S.A. and Shirô HATTORI, eds

- 1983 *Language atlas of the Pacific area, part 2: Japan area, Taiwan, Philippines, mainland and insular south-east Asia.* Canberra: The Australian Academy of the Humanities, in collaboration with the Japan Academy. PL, C-67.

YAMADA, Yukihiro

- 1966 A preliminary Itbayaten vocabulary. Institute of Asian Studies, University of the Philippines, Quezon City. Mimeographed.

YAP, Fe Aldave

- 1977 *A comparative study of Philippine lexicons.* Manila: Institute of National Language, Department of Education and Culture.

ZORC, R. David

- 1972 The western subgroup of Bisayan. *Oceanic Linguistics* 11/2:110-139.

- 1974a Towards a definitive Philippine wordlist - the qualitative use of vocabulary in identifying and classifying languages. *Oceanic Linguistics* 13:409-455.

- 1974b Internal and external relationships of the Mangyan languages. *Oceanic Linguistics* 13:561-600.

- 1977 *The Bisayan dialects of the Philippines: subgrouping and reconstruction.* PL, C-44.

- 1982 Micro- and macro-subgrouping: criteria, problems, and procedures. In Rainer Carle et al, eds *Gava^C: studies in Austronesian languages and cultures, dedicated to Hans Kähler*, 305-320. Berlin: Dietrich Reimer.

- in progress *Core etymological dictionary of Filipino.* Manila: Linguistic Society of the Philippines. [Fascicles 1-3, "A-H" have been printed in LSP Publications Nos.12,13,14.]

- in press Review of Blust *Austronesian etymologies.* *Philippine Journal of Linguistics.*

The genetic relationships of Philippine languages have been a subject of considerable interest and controversy. The traditional view, based on the work of Blust (1975) and others, has been that the Philippine languages form a single, coherent family, with the Tagalog group as the most central and the Ilocano group as the most peripheral. This view has been challenged by a number of scholars, including Zorc (1986), who has argued that the Philippine languages are in fact a collection of unrelated languages, with the Tagalog group being a distinct, non-Indo-European language family.

Zorc's argument is based on a number of linguistic features, including the lack of cognates between the Tagalog group and other Philippine languages, and the presence of unique morphological and syntactic features in the Tagalog group. He also points to the fact that the Tagalog group is geographically isolated from other Philippine languages, and that it has a long history of contact with non-Indo-European languages, such as the Austronesian languages of the Pacific Islands.

Despite Zorc's arguments, the traditional view remains the dominant one in the field of Philippine linguistics. This is largely due to the work of Blust and others, who have provided a comprehensive and detailed account of the genetic relationships of the Philippine languages. Their work has been widely cited and has shaped the way in which the Philippine languages are understood and studied.

However, there is still a need for further research into the genetic relationships of the Philippine languages. This is particularly true in the case of the Tagalog group, which remains a mystery to many linguists. Further research into the Tagalog group, and into the other Philippine languages, may help to clarify the genetic relationships of these languages and to provide a more complete picture of the linguistic history of the Philippines.

In conclusion, the genetic relationships of the Philippine languages are a complex and controversial issue. While the traditional view remains the dominant one, there is a growing body of evidence that challenges this view. Further research into the Tagalog group and other Philippine languages is needed to clarify the genetic relationships of these languages and to provide a more complete picture of the linguistic history of the Philippines.

The genetic relationships of the Philippine languages have been a subject of considerable interest and controversy. The traditional view, based on the work of Blust (1975) and others, has been that the Philippine languages form a single, coherent family, with the Tagalog group as the most central and the Ilocano group as the most peripheral. This view has been challenged by a number of scholars, including Zorc (1986), who has argued that the Philippine languages are in fact a collection of unrelated languages, with the Tagalog group being a distinct, non-Indo-European language family.

Zorc's argument is based on a number of linguistic features, including the lack of cognates between the Tagalog group and other Philippine languages, and the presence of unique morphological and syntactic features in the Tagalog group. He also points to the fact that the Tagalog group is geographically isolated from other Philippine languages, and that it has a long history of contact with non-Indo-European languages, such as the Austronesian languages of the Pacific Islands.

Despite Zorc's arguments, the traditional view remains the dominant one in the field of Philippine linguistics. This is largely due to the work of Blust and others, who have provided a comprehensive and detailed account of the genetic relationships of the Philippine languages. Their work has been widely cited and has shaped the way in which the Philippine languages are understood and studied.

However, there is still a need for further research into the genetic relationships of the Philippine languages. This is particularly true in the case of the Tagalog group, which remains a mystery to many linguists. Further research into the Tagalog group, and into the other Philippine languages, may help to clarify the genetic relationships of these languages and to provide a more complete picture of the linguistic history of the Philippines.

In conclusion, the genetic relationships of the Philippine languages are a complex and controversial issue. While the traditional view remains the dominant one, there is a growing body of evidence that challenges this view. Further research into the Tagalog group and other Philippine languages is needed to clarify the genetic relationships of these languages and to provide a more complete picture of the linguistic history of the Philippines.

The genetic relationships of the Philippine languages have been a subject of considerable interest and controversy. The traditional view, based on the work of Blust (1975) and others, has been that the Philippine languages form a single, coherent family, with the Tagalog group as the most central and the Ilocano group as the most peripheral. This view has been challenged by a number of scholars, including Zorc (1986), who has argued that the Philippine languages are in fact a collection of unrelated languages, with the Tagalog group being a distinct, non-Indo-European language family.

Zorc's argument is based on a number of linguistic features, including the lack of cognates between the Tagalog group and other Philippine languages, and the presence of unique morphological and syntactic features in the Tagalog group. He also points to the fact that the Tagalog group is geographically isolated from other Philippine languages, and that it has a long history of contact with non-Indo-European languages, such as the Austronesian languages of the Pacific Islands.

Despite Zorc's arguments, the traditional view remains the dominant one in the field of Philippine linguistics. This is largely due to the work of Blust and others, who have provided a comprehensive and detailed account of the genetic relationships of the Philippine languages. Their work has been widely cited and has shaped the way in which the Philippine languages are understood and studied.

A GENETIC GROUPING OF OCEANIC LANGUAGES IN BOUGAINVILLE AND THE WESTERN SOLOMONS

Malcolm Ross

1. INTRODUCTION

We take as our starting point for this paper¹ the establishment by Tryon and Hackman (1983:56-64) of a Western Solomons grouping of Oceanic languages, a grouping which includes the Oceanic languages of the Shortlands, Choiseul, the New Georgia group and Santa Ysabel (except Mbughotu) (see Map 1).

The eastern boundary of this group is clearly defined by the fact that the Western Solomons languages have merged Proto-Oceanic *d and *R, but (except Mono-Alu of the Shortlands) reflect Proto-Oceanic *l as a separate phoneme, whilst the languages of the South-East Solomons grouping (Pawley 1972; Tryon and Hackman 1983:65-66) have merged Proto-Oceanic *l and *R, but reflect Proto-Oceanic *d as a separate phoneme.

The western boundary of the Western Solomons grouping is not indicated by Tryon and Hackman, as their survey stops at the political boundary of the Solomon Islands. The westernmost language of the Western Solomons grouping, however, Mono-Alu of the Shortland Islands, clearly belongs to the Bougainville group (Ross 1982), whilst the major shared innovation of the Western Solomons languages, the merger of *d and *R, is also common to many of the Oceanic languages of Papua New Guinea (Milke 1965). The Western Solomons grouping is therefore open-ended at its western end, and the purpose of this paper is to explore its relationships with languages of the islands to its north-west, namely Bougainville and its outliers. We propose that the Oceanic languages of Tryon and Hackman's Western Solomons grouping, i.e. of the Shortlands, Choiseul, the New Georgia group and Santa Ysabel, and those of Bougainville, Buka and Nissan (Green) Islands form a single grouping, which we will call North-West Solomonic.²

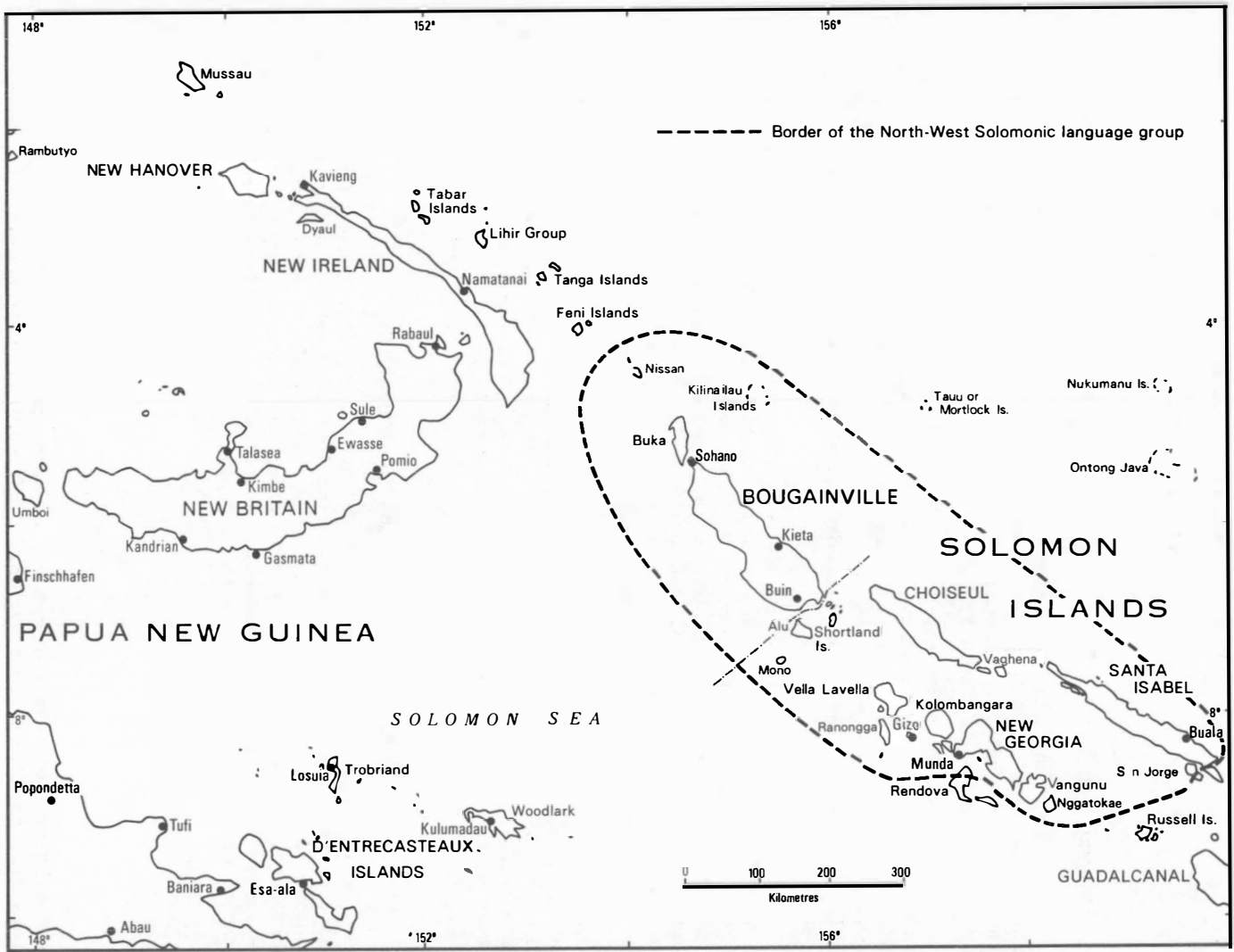
2. THE NORTH-WEST SOLOMONIC GROUP: SHARED INNOVATIONS

The languages of the North-West Solomons group³ are shown on Map 2 and are listed in Figure 2-1. Information on the location, naming, and internal sub-grouping of these languages is given for Bougainville by Ross (1982) and for the Shortlands, Choiseul, the New Georgia group and Ysabel by Tryon and Hackman (1983). The languages of this group reflect the following innovations (a table of sound correspondences is contained in the appendix):⁴

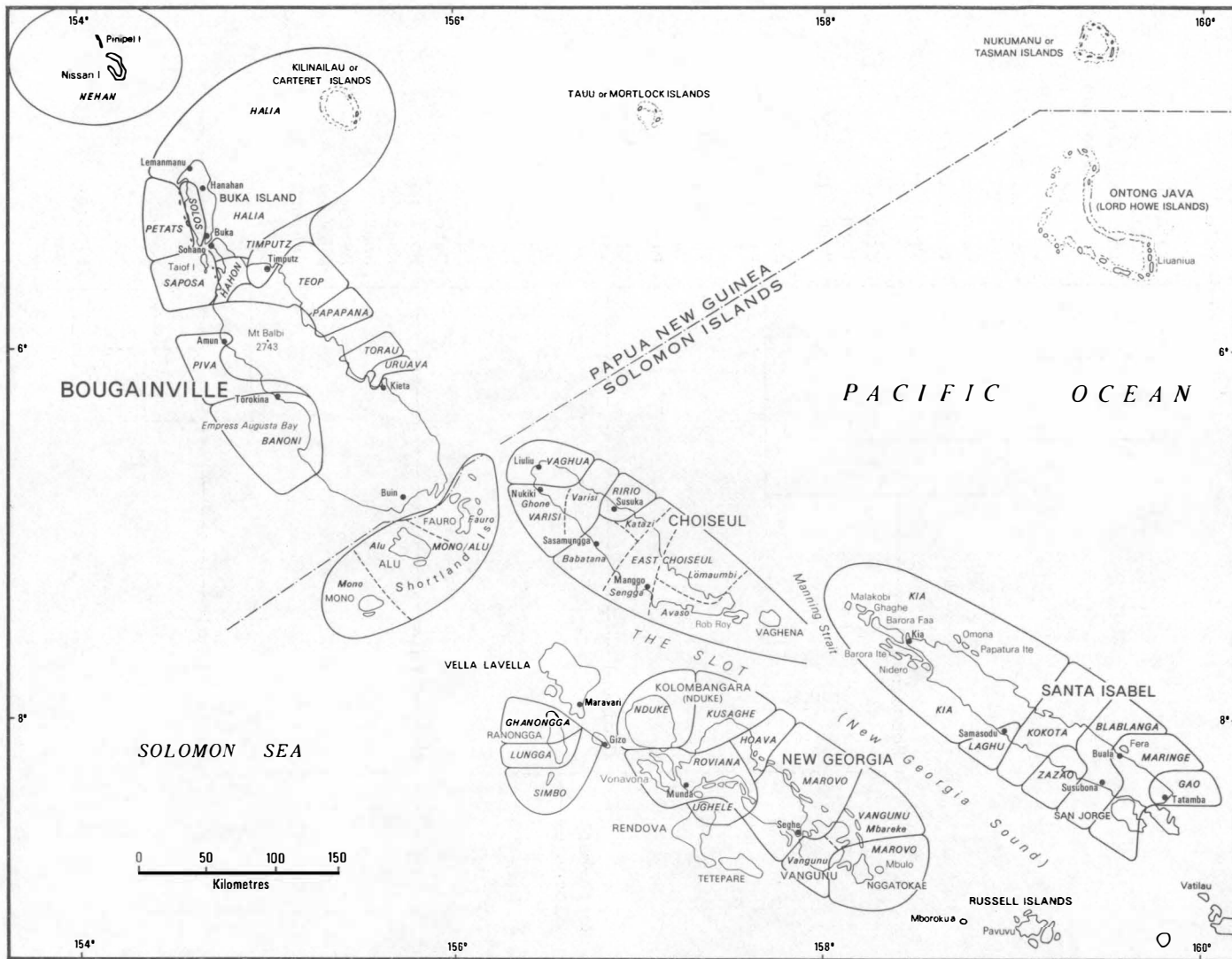
Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 175-200. *Pacific Linguistics*, C-94, 1986.

© Malcolm Ross

175



Map 1: Location of the North-West Solomonic group of Oceanic languages



Map 2: Oceanic languages of the North-West Solomonic group

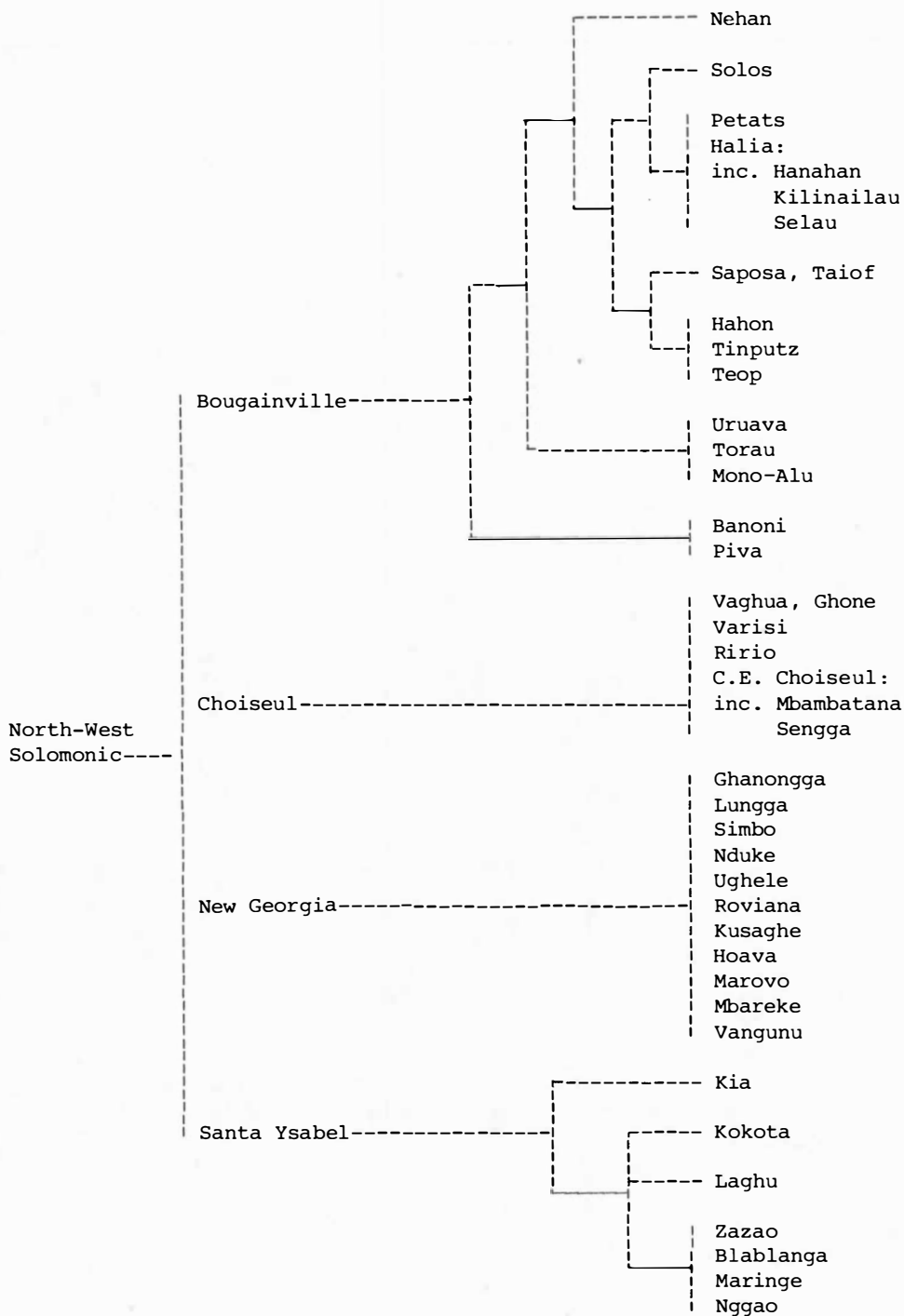


Figure 2-1: Oceanic languages of the North-West Solomons

- A. !OC *d and *R have merged as !NS *r;
- B. !OC *nd and *nt have merged as !NS *d;
- C. !OC *s and *j have merged as !NS *s;
- D. !OC *k has split into !NS *k and !NS *g;
- E. !OC *p has split into !NS *p and !NS *v;
- F. !OC word-final *-q has become the stop !NS *-k (whereas !OC initial and medial *q has merged with !NS *g (< !OC *k) or has been lost).
- G. !OC *w has been lost.
- H. All languages except Ririo reflect a !NS echo vowel added after !OC word-final consonants.
- I. Languages (except Nehan) which retain reflexes of the non-third person !OC disjunctive pronouns have an accreted !NS *r- before the initial vowel of each earlier vowel-initial pronoun.

Each of these innovations⁵ is briefly discussed and, where necessary, illustrated below.

2.1 The merger of !OC *d and *R

Innovation A, the merger of !OC *d and *R as !NS *r is illustrated in the examples below:⁶

- (1) !OC *dani *daytime* > !NS *rani

Bougainville: Nehan (ma-)rin, Solos nan, Petats len, Haku lan, Selau ran, Taiof naiŋ, Banoni nam, Uruava rani, Torau rare, Mono-Alu lale

Choiseul: Vaghua ranə, Varisi rani, Ririo ren, Mbambatana rani, Sengga rani

New Georgia: Lungga, Nduke, Roviana, Hoava, Vangunu rane

Ysabel: Kokota nare^a, Ghove, Maringe (na-)rane

^aProbable metathesis.

- (2) !OC *dua *two* > !NS *rua

Bougainville: Nehan (to-)ri(-ki), Solos nu, Petats (hu-)lu, Haku (to-)l, Selau (i-)ra, Taiof (fua-)n, Papapana nu⁷a(-ta), Banoni (too-)m, Piva (to-)nua, Uruava rua, Torau (a-)rua, Mono-Alu (e-)lua

Choiseul: Vaghua (ka-)rua, Varisi (ka-)rua, Ririo (ke-)r, Mbambatana (ke-)re, Sengga (ke-)ri

New Georgia: Lungga (ka-)ru, Nduke (ko-)ri, Roviana, Hoava, Vangunu (ka-)rua

Ysabel: Kia (pa-)lu(-gu), Kokota, Laghu (pa-)lu

- (3) !OC *tuqud *stand* > !NS *tuquru

Bougainville: Nehan, Selau tur, Taiof tu-tun, Hahon cun, Tinputz, Teop sun, Papapana toonu, Banoni cigom, Piva cugonu

New Georgia: Lungga, Nduke turu

- (4) !OC *Ruma(q)
- house*
- > !NS *ruma

Bougainville: Solos, Taiof, Banoni, Mono-Alu numa, Petats, Haku luma, Selau, Uruava, Torau ruma

New Georgia: Lungga ruma, Nduke ru-ruma *chest*

Ysabel: Maringe ruma *inside*

- (5) !OC *loŋoR
- hear*
- > !NS *loŋoro

Bougainville: Nehan loŋor, Solos, Taiof noŋon, Petats loŋon, Hako loŋolo, Selau roŋono, Papapana, Banoni, Piva noŋono, Uruava (ba-)ror(-i), Mono-Alu nono

Choiseul: Vaghua ŋoŋoro, Lungga noŋoro

New Georgia: Lungga noŋoro

- (6) !OC *[qa]paRa
- shoulder*
- > !NS *vara

Bougainville: Haku hala-hala, Selau war-wara, Torau ara, Mono-Alu hala

New Georgia: Lungga, Nduke, Roviana vara

Ysabel: Laghu fara

2.2 The merger of !OC *nd and *nt

Innovation B, the merger of !OC *nd and *nt as !NS *d rests on an interpretation of the data which differs somewhat from Tryon and Hackman's (1983:56). They found that items in their data corpus which are usually reconstructible with !OC *nd (e.g. !OC *ndaun *leaf*) reflect this apparent *nd in the same way as !OC *d and *R, and they therefore took it that !OC *nd had merged with *d and *R in Western Solomons languages. There is, however, in our corpus, a small number of items showing a correspondence set which requires the reconstruction of !NS *d, evidently reflecting !OC *nd and *nt (as distinct from !NS *r, which reflects !OC *d and *R). We take it that Tryon and Hackman's putative *nd items reflect !OC oral-grade *d (i.e. !OC *daun *leaf*). Examples of !NS *d are:

- (7) !OC *mu[n]di
- back*
- > !NS *mudi

Bougainville: Nehan mudi, Petats, Haku muru, Selau, Banoni muri, Tinputz puri^a, Teop buri^a, Mono-Alu muri-muri *later*

New Georgia: Nduke, Roviana mudi

^aThese forms and Proto-South-East Solomonian *buri reflect a !OC doublet *mpu[n]di.

- (8) !OC *mandindiŋ
- cold*
- > !NS *madidiŋ

Bougainville: Haku maririŋ

- (9) !OC *kandoRa
- phalanger*
- > !NS *kadora

New Georgia: Lungga gadora

Ysabel: Maringe k^hadara

- (10) !OC *punti *banana* > !NS *vudi

Bougainville:^a Solos hut, Petats hur, Selau wur, Taiof fur, Teop vuri, Uruava vudi, Torau udi

New Georgia: Lungga vudi

^aAll Bougainville reflexes mean (*banana*) *ripe*⁷

- (11) !OC *-nti[a] *their* > !NS *-di[a]

Bougainville: Solos, Petats, Teop, Banoni -ri, Haku -re, Selau, Taiof -r, Piva, Mono-Alu -ria, Uruava -di, Torau -dia

New Georgia: Lungga -di, Nduke -di-ria, Roviana, Hoava -dia, Vangunu -di-eri-kia

Ysabel: Kia, Kokota, Laghu, Blablanga, Ghove, Maringe -di

2.3 The merger of !OC *s and *j

Innovation C, the merger of !OC *s and *j, is common to much of Oceania (but has not occurred in certain Admiralty Islands languages, where Proto-Austronesian *j is retained as a separate phoneme (Blust 1978), nor in some Fijian communalects (Geraghty, this volume)). Since the value of this merger as a subgrouping criterion is consequently low, it is not illustrated here.

2.4 The split of !OC *k

The important feature of Innovation D, the split of !OC *k into !NS *k and *g, is that, for the large majority of items containing !OC *k, daughter-languages in the North-West Solomonian region agree in reflecting either !NS *k or !NS *g in a given item - we do not encounter a haphazard mixture of *k- and *g-reflexes. This is illustrated below:

!OC *k- > !NS *k-:

- (12) !OC *k[a,u](n) *supe rat* > !NS *ku(s,j)u^a

Bougainville: Nehan kih, Solos, Sealau, Taiof, Banoni kiso, Hahon kuso, Teop kuho

Choiseul: Vaghua kōj, Varisi kuzu, Ririo kuj

Ysabel: Maringe na-kusi

^aUnexpected loss of !OC -pe.

- (13) !OC *kai *one* > !NS *kai, *kai-sa

Bougainville: Uruava ka-a, Torau ka-sa, Mono-Alu ka-la

Choiseul: Vaghua, Varisi ka-la, Ririo ki-ki, Mbambatana kə-ke, Sengga kō-ke

New Georgia: Nduke ke-ka, Roviana, Hoava ke-ke, Vangunu me-ka

Ysabel: Kia kai-ke-gu, Kokota kai-ke-u, Laghu kai-ke, Blablanga kai-sa, Ghove, Maringe ka-ha

!OC *-k- > !NS *-k-:

(14) !OC *mpaku *cheek* > !NS *baku

Bougainville head: Solos bak, Petats, Haku baku

Ysabel: Kia bako

!OC *-k- > !NS *g-:

(15) !OC *kani *eat* > !NS *gani

Bougainville: Nehan, Tinputz en, Solos h-en, Petats n-en, Haku n-an, Taiof aiŋ, Hahon an-an, Teop an, Papapana, Torau ani, Uruava ana, Mono-Alu aaŋ

New Georgia: Lungga ga-gani, Roviana gani-gani

Ysabel: Kia gani-ni

(16) !OC *kutu *louse* > !NS *gutu

Bougainville: Nehan ut, Solos pe-h-ut, Petats pi-h-ut, Haku wutu, Selau wut, Teop us, Papapana ut-auno, Piva gucu, Uruava, Mono-Alu utu

Choiseul: Vaghua gec, Varisi utu, Ririo v-uc, Mbambatana v-utu, Sengga v-otu

New Georgia: Nduke v-utu, Roviana, Hoava gutu, Vanguu utu

Ysabel: Kia, Laghu gutu, Kokota, Blablanga gutu

!OC *-k- > !NS*-g-:

(17) !OC *ikan *fish* > !NS *igana

Bougainville: Nehan, Taiof ian, Solos iean, Petats, Tinputz ien, Haku, Hahon, Teop, Papapana, Uruava, Mono-Alu iana, Selau iena, Piva v-igana, Torau iala

Choiseul: Vaghua, Varisi iŋana^a

New Georgia: Lungga, Nduke, Roviana, Hoava igana, Vanguu ihana

^aŋ for expected g

(18) !OC *matakut *fear* > !NS *matagutu

Bougainville: Solos ratout^a, Petats matout, Haku matutu, Selau matut

New Georgia: Lungga matagu, Roviana matagutu

Ysabel: Blablanga, Ghove, Maringe mhagu

^ar- unexplained.

In North-West Solomonic, !OC *k is far more frequently reflected as !NS *g than !NS *k, and of the items reflecting !NS *k, the large majority are instances of initial *k. This is consistent with the hypothesis that, at some stage prior to North-West Solomonic, a process changing !OC *k to *g had got under way, but been arrested before it had spread to all potential candidates in the lexicon. Since it is predictable that this change would affect medial *k before initial *k, the pattern of the data is consistent with the hypothesis.

It could well be argued that the presence of *k items is due to borrowing, and there is indeed a small sprinkling of items in which languages do not agree in their reflex of !OC *k. Thus in the case of !OC *mpakiwa *shark* below, it is

probable that the reflexes in example (19) reflect an early borrowing, since the *-u- of !NS *baku-a-i possibly reflects !OC *-w-, which is normally lost in Proto-North-West Solomonic, whilst the reflexes in example (20) are expected forms:

(19) !OC *mpakiwa *shark* > !NS *baku-a-i

Bougainville: Nehan bakue, Solos bake, Haku baki, Teop, Uruava baku-baku, Banoni bakuo, Mono-Alu baʔoi

Choiseul: Vaghua bakui, Varisi bakuai, Ririo boʔove

(20) !OC *mpakiwa *shark* > !NS *bagea

New Georgia: Lungga bagea

Ysabel: Kia, Kokota, Maringe bae-su

However, the distribution of the two reflexes of !OC *k and the fact that a similar distribution is found both to the north-west (in New Ireland) and to the south-east (in South-East Solomonic; cf. Levy n.d.) makes it unlikely that borrowing has played any recent major role in producing two reflexes of !OC *k. Indeed, to the north-west, k- and ɣ-reflexes occur in much the same items as they do in the North-West Solomonic languages.

2.5 The split of !OC *p

Innovation E, the split of !OC *p into !NS *p and *v, is the labial equivalent of innovation D, and is illustrated below:

!OC *p- > !NS *p-:

(21) !OC *posi *squeeze* > !NS *po(z,j)i

Bougainville: Nehan pos, Solos pot, Petats, Haku, Selau, Taiof poc, Teop pos

Ysabel: Ghove boji, Maringe poji

(22) !OC *pusi *break wind* > !NS *pusi

Bougainville: Petats pus *buttocks*, Haku pisi *break wind*, Tinputz pih, Teop pih

Choiseul: Mbambatana pusu

New Georgia: Roviana pisi

(23) !OC *patu *round object* > !NS *patu

Bougainville head: Taiof, Papapana, Uruava patu, Hahon pac, Tinputz, Teop pasu, Torau pau

Choiseul seed:^a Vaghua pacə, Varisi patu-ru, Ririo puce, Mbambatana pōti, Sengga pətu

New Georgia seed:^a Vangunū patu

Ysabel head: Kia, Kokota pau, Blablanga, Ghove, Maringe p^hau

^aA nasal-grade form, !OC *mpatu, !NS *batu, is reflected in Choiseul with the meaning *head* or *chieftain*, and in New Georgia with the meaning *head*.

(24) !OC *pujaq *foam* > !NS *puzaka

Bougainville: Haku posa, Teop poha

Choiseul: Mbambatana pusaka

New Georgia: Roviana puzaka

!OC *p- > !NS *v-:

(see also example (6))

(25) !OC *pitu *seven* > !NS *vitu

Bougainville: Nehan (to-)witi, Solos hit, Petats (to-)hiet, Haku (to-)hiti, Taiof fit, Uruava uitu, Torau tu, Mono-Alu hitu

Choiseul: Vaghua (ka-)vuc, Varisi (ka-)vitu, Ririo z-iuc, Mbambatana vitu, Sengga vettu

New Georgia: Lungga, Nduke vitu

Ysabel: Kia vitu(-gu), Kokota fitu(-gu), Laghu, Blablanga, Ghome, Maringe fitu

(26) !OC *papine *woman* > !NS *vavine

Bougainville: Petats hihin *woman*, Haku hahine *cross-sibling*, Taiof fafine *cross-sibling*, Hahon, Papapana vavine *cross-sibling*, Torau baina^a *woman*, Mono-Alu hahine *cross-sibling*

Choiseul cross-sibling: Vaghua vavene, Varisi vavani, Ririo vavivi, Mbambatana, Sengga vavini-a

New Georgia: Roviana vine(-ki)^b *female*, Hoava hina(-guru)^b *woman*, Vangunu vavene *sister*

^ab- for expected Ø. ^b!OC *pa- lost.

!OC *-p- > !NS *-v-:

(see also !OC *papine above)

(27) !OC *sapa *what* > !NS *sava

Bougainville: Nehan hawa, Solos sah, Petats, Taiof, Torau sa, Haku ha, Selau (n-)awa, Banoni sava, Uruava (n-)ava, Mono-Alu aha

Choiseul: Vaghu ava(-na), Varisi, Ririo, Mbambatana ava

New Georgia: Lungga sa, Nduke saga, Roviana, Hoava (na-)sa

(28) !OC *lipo(n) *tooth* > !NS *livo

Bougainville: Nehan liwo, Solos, Mono-Alu niho, Petats, Haku liho, Selau luwo, Tinputz livo, Teop rivo, Torau nio(-tau)

New Georgia: Lungga, Nduke, Roviana livo, Vangunu livo(-no)

The distribution of !NS *p and *v reflexes of !OC *p resembles that of !NS *k and *g reflexes of !OC *k, except for the fact that !NS *p never occurs medially in reflexes of Proto-Oceanic items, indicating that the process which changed *p to *v was completed, or nearly so, in items with medial *p.

2.6 The reflex of !OC word-final *-q

Innovation F is that !OC word-final *-q became the stop !NS *-k. Sound correspondences relevant to this innovation are set out in Table 2-1.

Table 2-1: North-West Solomons reflexes of Proto-Oceanic *k and *q

!OC	*k ₁	*k ₂	*q	*q/_#
!NS	*k	*g	*q	*k
Nehan	.	∅	∅	∅
Solos	k	∅	∅	.
Petats	k	∅	∅	.
Haku	k	∅	∅	∅
Selau	k	∅	∅	k
Taiof	k	∅	∅	k
Hahon	k	∅	∅	.
Tinputz	k	∅	∅	∅
Teop	k	∅	∅	∅
Papapana	.	∅	∅	∅
Banoni	k	g-; -g-, -∅-	g, ∅	g
Piva	k	g-; -g-, -∅-	∅-; -g-, -∅-	.
Uruava	k-	∅	∅	∅
Torau	k-, ∅-	∅	∅	∅, k
Mono-Alu	k, ?-, ∅-	∅	∅	?
Vaghua	k	g	g, ∅	k
Varisi	k	g, ∅	g, ∅	k
Ririo	k, ?	g, ∅	∅	?
Mbambatana	k	g, ∅	∅	.
Sengga	k	∅	∅	.
Lungga	k	g-, ∅-; -g-	∅	k
Nduke	k	g, ∅	∅, g	.
Roviana	k	g	∅	k
Hoava	k	g	∅	k
Vangunu	k	∅, h, g	∅	.
Kia	k	g	g-; -∅-	k
Kokota	k	g, g; g/_C	g-; -∅-	.
Laghu	k	g	g-; -∅-, -g-	k
Blablanga	k, kh	g, g	g	.
Ghove	k, kh	g; g/_C	g-, ∅-; -∅-	.
Maringe	k, kh	g; g/_C	g-, ∅-; -∅-	.

The columns headed !OC *k₁ and !OC *k₂ represent the two correspondence sets discussed under innovation D above, and giving rise to the reconstruction of !NS *k and *g. !OC initial and medial *q have been deleted or merged with !NS *g in North-West Solomonic languages. For example (see also example (3), !OC *tuqud *stand*):

(29) !OC *mpaReqo *breadfruit* > !NS *bareqo

Bougainville: Nehan, Selau, Torau bario, Haku, Mono-Alu baleo, Taiof vare, Tinputz panio, Teop baneo, Uruava bareo

Choiseul: Vaghua, Varisi barego, Ririo, Mbambatana, Sengga bario

Ysabel: Blablanga, Ghove, Maringe hnego

!OC final *-q, however, shows a stop reflex in a number of languages, suggesting that the !NS reflex of !OC *-q was a voiceless velar stop. For example (see also example (24), !OC *pujaq *foam*):

- (30) !OC *mpuaq *areca nut* > !NS *buaka

Bougainville: Selau boko, Taiof bok, Banoni bugava (metathesis), Torau buka

- (31) !OC *toloq *eel* > !NS *toloko

Choiseul: Vaghua, Varisi toloko

New Georgia: Hoava toloko

- (32) !OC *papaq *short* > !NS *papaka

Bougainville: Mono-Alu papa?a-na

New Georgia: Lungga, Roviana papaka

- (33) !OC *luaq *tears* > !NS *luak(a)

Choiseul: Ririo lua? *water*

Of particular note is the fact that !OC final *-k is *not* preserved as a stop, but, having acquired an echo vowel in Proto-North-West Solomonic, has behaved like !OC medial *k₂, and become !NS *g:

- (34) !OC *tasik *salt water* > !NS *tasigi

Bougainville: Selau, Taiof tasi, Banoni tagisa (metathesis)

Ysabel: Kia, Kokota, Blablanga tahi

- (35) !OC *manuk *bird* > !NS *manugu

Bougainville: Banoni, Piva manugu

New Georgia: Lungga manugu, Roviana manuvu, Vangunu maunu

- (36) !OC *mpoRok *pig* > !NS *borogo

Bougainville: Papapana boro, Banoni borogo

New Georgia: Lungga borogo, Nduke bogoro (metathesis)

The correspondences in Table 2-1 suggest that !OC *k first underwent the process which turned the majority of its occurrences into !NS *g. After this had occurred, !OC non-final *q merged with *g or was lost, but !OC final *q became a voiceless velar stop.

2.7 The loss of !OC *w

Innovation G, the loss of !OC *w, is illustrated in example (20) above and in the following:

- (37) !OC *siwa *nine* > !NS *sia

Bougainville: Nehan (lu-)sio, Solos sie, Petats, Selau (to-)sie, Haku (to-)si, Taiof, Torau sia, Banoni v-isa^a, Uruava ia, Mono-Alu (u-)lia

Choiseul: Vaghua (ka-)x-əsa (metathesis), Varisi ka-ia, Ririo, Mbambatana, Sengga zia

New Georgia: Lungga, Nduke, Roviana, Hoava, Vangunua sia

^a!NS *sia > Pre-Banoni *isa (by metathesis) > Banoni v-isa (by regular v-accretion)

- (38) !OC *lawe *plumage* > !NS *lae

Bougainville: Nehan la-le

Ysabel: Kia lae *wing*

- (39) !OC *walu *eight* > !NS *alu

Bougainville: Nehan (to-)ali, Solos (to-)an, Petats, Haku, Selau (to-)wal, Taiof (ji-)an, Uruava aru, Torau anu, Mono-Alu alu

Choiseul: Vaghua (ka-)z-al, Varisi (ka-)z-alu, Ririo z-ol

- (40) !OC *wakaR *root* > !NS *ağara

Bougainville: Selau ara, Hahon, Teop ana, Tinputz an, Uruava, Torau ağara

New Georgia: Lungga, Nduke ağara, Hoava ağoro

Ysabel: Kia z-ağara, Kokota, Blablanga z-agra

- (41) !OC *waRoj *string* > !NS *aroso

Bougainville: Mono-Alu alolo

Ysabel: Kia (n-)aroho, Kokota (n-)arho, Maringe (ñ-)arho

- (42) !OC *kawil *fish hook* > !NS *ğaili

Bougainville: Nehan i-ail, Solos enani^a, Petats ilali^a, Haku lali, Taiof ir, Teop iri, Banoni airi, Torau aini, Mono-Alu aili

New Georgia: Lungga ğaili

Ysabel: Kokota ğaili

^aApparently !NS ğaili > Pre-Proto-Buka *aili (expected form) > Proto-Buka *ilali (by metathesis and reduplication)

This interpretation of the data is a little different from Tryon and Hackman's (1983:61), who find two instances in which !OC *w appears to be reflected by f or v in Ysabel languages and therefore suggest that !OC *w has not been lost on Ysabel. However, examples (20), (38), (40), (41), and (42) above indicate loss of !OC *w on Ysabel. The examples cited by Tryon and Hackman are:

(43) !OC *siwa *nine* >*Ysabel*: Kokota hnevau, Blablanga hneva, Ghove, Maringe heva(44) !OC *ma-ñawa *breathe* >*Ysabel heart*: Kokota, Blablanga na-nafa, Ghove ña-ñafa, Maringe na-ñafa

In the case of example (43), it is probable that the forms cited are not derived (or not directly derived) from !OC *siwa. The expected *Ysabel* reflex of !OC *siwa, !NS *sia is *hia, and this is indeed found in Nggao fa-hia. Since Nggao has otherwise expected reflexes of the numerals, prefixing those from *five* upwards with fa-, it is reasonable to suppose that Nggao reflects the directly inherited form of !OC *siwa and that the forms in example (43) have some other source, about which, however, we can only speculate. The forms in example (44) remain unexplained. Ghove and Maringe reflect expected ñ, but the -f- of -ñafa suggests !NS *-ñava, not expected *-ñawa. However, Vaghua (Choiseul) ma-nava *liver* also reflects !NS *-ñava, indicating that an idiosyncratic innovation may have taken place in this item.

2.8 Echo vowels

Innovation H, whereby an echo vowel was added in Proto-North-West Solomonic after a Proto-Oceanic final consonant (which vowel is subsequently lost in a minority of daughter-languages under the same conditions as they lose Proto-Oceanic final vowels) is liberally illustrated in examples throughout this paper and is not further discussed here.

2.9 Accreted *r- on pronouns

Innovation I states that languages (except Nehan) which retain reflexes of the non-third person !OC disjunctive pronouns have an accreted !NS *r- before each earlier vowel-initial pronoun. This statement takes as its starting point two of Tryon and Hackman's observed innovations, namely that the languages of Choiseul, the New Georgia group and Santa *Ysabel* all reflect a form *a-rau I, instead of the expected *(n)au (1983:57), and that Choiseul reflexes of !OC *koe *thou*, *kami *we* (excl.) and *kamu *you* all reflect initial *r- instead of initial *k- (1983:60). However, the relevant data from Bougainville give grounds for combining and extending these two observations.

A few languages have non-third person disjunctive pronoun forms which cannot be attributed to the disjunctive pronouns of Proto-Oceanic or Proto-North-West Solomonic, and these will not be considered here.⁸ Forms with which we are concerned are listed in Table 2-2.

A large number of the pronouns in Table 2-2 show a prefixed e-, a-, ia- or o-, and in the North Bougainville languages (Solos to Teop in Table 2-2) each pronoun has a non-focal variant from which the prefix is missing. There is good reason to reconstruct three of these prefixes as articles in Proto-North-West Solomonic: *a 'common singular', *o 'common mass/plural', *e 'personal'. Their distribution in Table 2-2 suggests that *o was not normally used with pronouns, but that *e and *a certainly were, although the usage of *a 'common article' (instead of expected *e 'personal article') remains unexplained. Torau -di and the suffixes on *Ysabel* plural pronouns are number-markers, whilst the suffixes on Choiseul plural pronouns are subject pronouns, originally proclitics to the verb, which have become enclitics to the disjunctive pronoun.

Table 2-2: Selected disjunctive pronoun forms in North-West Solomonian languages

	<i>I</i>	<i>thou</i>	<i>we</i> (excl.)	<i>you</i>
Nehan	+ [i]o	+ [i]a	+ [i]em ^a	+ om ^a
Solos	[e]na	[e]no	+ [e]mem	+ [e]mu
Petats	e-lia	e-lou	e-lam	e-lomi
Haku	.	a-lo	a-lam	a-limu
Hanahan	a-lia	a-lo	a-lam	a-limiu
Selau	a-la	a-li	+ a-mam	+ a-mu
Taiof	a-iña	a-noh	+ a-mam	a-ñim
Hahon	e-ne	.	+ e-mam	+ e-am
Tinputz	e-iö	.	+ e-möm	+ e-öm
Teop	e-na	.	e-nam	+ e-am
Papapana	a-niau	a-nioi	.	+ a-mu
Banoni	na	no	+ gamam	+ gam
Piva	a-na	+ a-go	+ a-gamam	+ a-gam[i]
Uruava	a-ria	a-ro	a-ramani	a-ramu
Torau	+ i-nau	+ i-ne	+ ni-mani-di	+ ni-mu[-di]
Vaghua	a-ra	+ a-go	o-re	o-ram
Varisi	e-ra	a-ro	remu-mi	ramu-mamu
Ririo	ra	ge-r	rem	ram
Mbambatana	ra	re	rami-mami	ramu-mamu
Sengga	a-ro	re	rami-mami	ramu
Lungga	a-ra	a-o	a-gei	a-gou
Nduke	rai	goi	gami	gamu
Roviana	a-rau	a-go	gami	gamu-kasa
Hoava	rao	goe	gami	gamu
Vangununu	e-ra	i-o	ami	amu-kia
Kia	a-ra	a-go	gai	gau
Kokota	a-ra	a-go	gai	gau
Laghu	a-ra	a-go	gai-har	gau-haro
Blablanga	a-ra	a-go	gai	go-tilo
Ghove	ia-ra	ia-go	ga-hati	go-tilo
Maringe	ia-ra	ia-go	ge-hatihui	go-tilo

^aTodd (1978:1184) gives [i]am for both *we* (excl.) and *you*. The forms used here are from my fieldnotes (my informant is from Yatchibol village, Nissan Island).

Note: The significance of + and of the box is explained in the text.

Pronouns which do not reflect accreted !NS *r- are marked in two ways in Table 2-2: the New Georgia and Ysabel pronouns other than *I* are boxed in, and the various occurrences of pronouns without accretion in Bougainville and Choiseul are indicated with a '+'. It is clear that many of these can be explained as reflexes of well-established Proto-Oceanic forms, with or without a reflex of a preceding article:⁹

(45)	!OC	!NS
<i>I</i>	*[i]au	*[i]au
<i>thou</i>	*ko[e]	*go[e]
<i>we</i> (excl.)	*kami	*gami
<i>you</i>	*kam[i]u	*gam[i]u

The most economic explanation of the *you*-forms Solos [e]mu, Selau, Papapana a-mu, Hahon, Teop e-am, Tinputz e-öm is that the expected !NS form *gamu, preceded by one or other of the articles (*a-gamu, *e-gamu), was reduced by expected loss of !NS *g to *a-amu, *e-amu. If the same sequence of decay is attributed to the *we* (excl.)-forms Solos [e]mem, Selau, Taiof a-mam, Hahon e-mam, Tinputz e-möm, then we need to reconstruct as an alternant to !NS *gami *we* (excl.) a !NS form *gamami (> *a-gamami, *e-gamami > *a-amami, *e-amami), which is in any case reflected in Banoni *gamam and Piva a-gamam.

We are left with a small number of *r-less forms not thus explained. These are Nehan [i]a *thou*, which will remain unexplained, and the Torau forms, which show a similarity to the forms of their Bougainville neighbours, but contain a morpheme n[i] whose origin is unclear.

The forms in Table 2-2 which *do* reflect accreted initial *r- have a skewed distribution: *I* always receives an accretion (except in Nehan, whose pronouns never receive it), whilst the other three pronouns receive it only in Bougainville and Choiseul and only sporadically. *I* apparently always receives accreted *r- because !OC/!NS *[i]au is vowel-initial. The other three forms do not always receive it because they are not vowel-initial: !NS initial *g- intervenes. It is therefore tempting to explain the sporadic occurrence of *r- as an accretion after the loss of !NS *g-, i.e., as an independent parallel innovation in different groups of languages with the occurrence of a new initial vowel. Several facts speak against this interpretation, however:

- a) *r- is almost universally present in *I*-forms. If we assume that *r-accretion occurred in the other forms after loss of *g-, then we must also assume not only that accretion after loss of *g- occurred independently in different groups of languages, but also that *r-accretion on *I*-forms similarly occurred independently, as it is part of the same process. Yet the near-universal presence of *r- on *I*-forms makes its reconstruction in Proto-North-West Solomonic a far more plausible hypothesis than independent parallel innovation. This in turn implies the reconstruction of forms with accreted *r- for all four pronouns, if we take it that all four underwent the same process.
- b) On the parallel innovation hypothesis, the Banoni *thou* form should be **go(i) (cf. Piva a-go(i)), as we do not expect loss of *g in Banoni, but in fact we find no; a similar argument applies to the *thou*, *we* (excl.) and *you* forms of the Choiseul languages, where, on the parallel innovation hypothesis, *g- has been inexplicably replaced by *r-. Both Banoni no *thou* and the Choiseul forms require the reconstruction of *r-forms in Proto-North-West Solomonic.
- c) If Lungga a-ra, Vangunu e-ra *I*, have acquired r- as the result of late independent innovation, then we would expect Lungga a-o *thou*, Vangunu i-o *thou*, ami *we* (excl.), and amu-kia *you* also to have acquired r-, but they have not. If, however, we take it that

Proto-North-West Solomonian had both *r-pronouns (whence the *I* forms) and *g-pronouns (whence a-o, i-o, ami and amu-kia with expected *g-loss), then the apparent anomaly is explained.

In view of these considerations we reconstruct vowel-initial alternant forms in Proto-North-West Solomonian, to which *r- was accreted before the break-up of the proto-language:

(46)	!OC	!NS			
			I	II	III
	<i>I</i>	*[i]au	*[i]au		*r-[i]au
	<i>thou</i>	*ko[e]	*go[e]	*o[e]	*r-o[e]
	<i>we(excl.)</i>	*kami	*gami	*ami	*r-ami
			*gamami	*amami	*r-amami
	<i>you</i>	*kam[i]u	*gam[i]u	*am[i]u	*r-am[i]u

This *r- has a probable cognate in the Nehan ligature -r-, which has a variety of functions, some now fossilised, some still productive. One of its productive functions is that it stands between a predicate and a following subject noun phrase:

- (47) a yan k-e eni-ni-r keketik
 ART fish ASP-SUBJ eat-OBJ-LIG child
 (It's) a fish the child is eating.

Such occurrences of -r- are common, as Nehan preposes a contrastive and/or newly introduced topic and postposes after -r- a subject noun phrase that does not meet one of these conditions. The ligature -r- is also used between a possessed noun phrase and its possessor noun phrase:

- (48) na tinihe-r toya
 POSS canoe-LIG chief
 The chief's canoe

In the North Bougainville languages the focal, preposed pronoun set is as shown in Table 2-2, whilst the postposed set, used as in Nehan, formally resembles the preposed set but lacks the prefixed article. Thus, for example, the preposed form Selau a-la *I* has a postposed equivalent -la:

- (49) ala e tara-ya tawor
I VI see-TR woman
I saw the woman
- (50) e Taga t-e tare-i-la
 ART Taga REL-VI see-OBJ-I
 (It was) Taga that I saw
- (51) a ruma ta-g-la
 ART house of-my-I
 My house.

We may perceive the role evidently played by the ligature *r- in the antecedents of the constructions in examples (50) and (51) as follows:

- (52) tare-i-la < *tara-i-r-au
see-OBJ-I see-OBJ:3s-LIG-I
- (53) ta-g-la < *ta-gu-r-au
of-my-LIG-I

It is likely that the ligature became a fossilised attachment to the disjunctive pronoun, generalised to all its uses.

2.10 Review

Nine innovations have been presented which are shared in common by the languages of the North-West Solomonic region. Whilst some of these innovations are also found outside the region and the value of the individual innovations as subgrouping criteria is varied, the fact that this group of around 35 languages shares this bundle of innovations is reasonable evidence that they have a common ancestor of a lower order than Proto-Oceanic.

APPENDIX

Bougainville and Western Solomons sound correspondences

!OC	*p ₁	*p ₂	*mp	*w	*m	*mw
!NS	*p	*v	*b	*∅	*m	.
Nehan	p	w; ∅/_u	b	∅	m	.
Solos	p	h	b	∅	m	.
Petats	p	h	b	∅	m	.
Haku	p	h	b	∅	m	m
Selau	p	w; -u	b	∅	m	.
Taiof	p	f	b, v	∅	m	.
Hahon	p	v	b	∅	m	.
Tinputz	p	v; -∅, -h	p	∅	m	.
Teop	p	v	b	∅	m	.
Papapana	p	v	b	.	m	.
Banoni	p	v; v, ʒ/#_o, u	b	∅	m	.
Piva	p	v	b	.	m	.
Uruava	p	v; ∅/_u	b	∅	m	.
Torau	p	∅	b	∅	m; -ŋ	.
Mono-Alu	p	h; -∅	b	∅	m; -ŋ	.
Vaghua	p	v	b	∅	m	.
Varisi	p	v	b	∅	m	.
Ririo	p	v; ∅/#_i, e	b	∅	m	.
Mbambatana	p	v	b	∅	m	.
Sengga	p	v	b	∅	m	.
Lungga	p	v	b	∅	m	.
Nduke	p	v	b	∅	m	.
Roviana	p	v	b	∅	m	.
Hoava	p	v	b	∅	m	.
Vangunu	p	v	b	∅	m	.
Kia	p	f	b	∅	m	m
Kokota	p; f/_C	f	b	∅	m	.
Laghu	p; f/_C	f; b/_C	b	.	m	m
Blablanga	ph; f/_C	f	b	∅	m	.
Ghove	ph; f/_C	f	b	.	m	.
Maringe	ph; f/_C	f	b	∅	m	.

APPENDIX (cont'd)

!OC	*t	*nt,*nd	*n	*ñ	*l
!NS	*t	*d	*n	*ñ	*l
Nehan	t	d	n	n	l
Solos	t	r; -t	n	n	n
Petats	t; c/ <u>i</u>	r	n	n	l
Haku	t; c/ <u>i</u>	r	n	n	l
Selau	t; c/ <u>i</u>	r	n	n	l,r
Taiof	t; c/ <u>i</u>	r	n; ŋ/ <u>*i</u> #	n	n,r
Hahon	t; c/ <u>i</u> ,u	.	n	.	n
Tinputz	t; s/ <u>i</u> ,u	r	n	.	n,l
Teop	t; s/ <u>i</u> ,u	r	n	.	n,r
Papapana	t; s/ <u>i</u>	r	n	.	n
Banoni	t; c/ <u>i</u> ,u	r	n; -m	n; -m	n,r
Piva	t; c/ <u>i</u> ,u	r	n	.	n,r
Uruava	t; s/ <u>i</u>	d	n	.	r
Torau	t	d	n	.	n,l
Mono-Alu	t	r	n; -ŋ; l/ <u>v</u> l,l <u>v</u>	.	l,n
Vaghua	t; c/ <u>i</u> ,u	.	n	n	l
Varisi	t; s/ <u>i</u>	.	n	.	l
Ririo	t; c/ <u>i</u> ,u	d	n	.	l
Mbambatana	t	d	n	.	l
Sengga	t	d	n	.	l
Lungga	t	d	n	ñ	l
Nduke	t	d	n	n	l
Roviana	t	d	n	n	l
Hoava	t	d	n	n	l
Vangunu	t	d	n	n	l
Kia	t	d	n	n	l
Kokota	t	d	n	n	l
Laghu	t	d	n	n	l
Blablanga	t,th	d	n	n	l; gl-
Ghove	t,th; k/ <u>C</u>	d	n	ñ	l; gl-
Maringe	t,th; k/ <u>C</u>	d	n	ñ	l; gl-

APPENDIX (cont'd)

!OC	*d, *R	*s, *j	*ns	*nj	*y	*k ₁
!NS	*r	*s	*z	*j	*∅	*k
Nehan	r	h	s	.	∅	.
Solos	n	s	s	r	.	k
Petats	l	s	s	r	∅	k
Haku	l	s	s	r	.	k
Selau	r, l	s	s	.	∅	k
Taiof	n; η/ *i; -r-, -n-	s	s	r	.	k
Hahon	n; -r-, -n-	s	.	.	.	k
Tinputz	n; -∅-	h	.	.	.	k
Teop	n; -n-, -r-	h	.	r	.	k
Papapana	n; -n-, -r-	s, t	.	r	.	.
Banoni	n; -n-, -r-; -m	s	s	.	∅-	k
Piva	n	s	.	.	∅	k
Uruava	r	∅	∅	d	∅	k-
Torau	r	s	s	d	∅	k-, ∅-
Mono-Alu	l, n	l, s	l	r	∅	k, ?-, ∅-
Vaghua	r	∅	s	j	.	k
Varisi	r	∅	s	z	.	k
Ririo	r	∅	s	j	.	k, ?
Mbambatana	r	∅	s	j	.	k
Sengga	r	∅	s	j	.	k
Lungga	r	s	∅	j	.	k
Nduke	r	s	∅, h	z	.	k
Roviana	r	s	h	z	.	k
Hoava	r	s	h	z	∅	k
Vangunu	r	s-; -z-	s, h	j	.	k
Kia	r, l	s-; -h-	h	z	∅	k
Kokota	r, l	s-; -h-	h	z	.	k
Laghu	r, l	s-; -h-	h	z	.	k
Blablanga	r, l	s-; -h-	h	z	.	k, kh
Ghove	r	s-; -h-	h	j	.	k, kh
Maringe	r, l	s-; -h-	h	j	.	k, kh

APPENDIX (cont'd)

!OC	*k ₂	*q	*q/_#	*ŋk	*ŋ
!NS	*g	*q	*k	*g	*ŋ
Nehan	∅	∅	∅	g	ŋ
Solos	∅	∅	.	g	ŋ
Petats	∅	∅	.	g	ŋ
Haku	∅	∅	∅	g	ŋ
Selau	∅	∅	k	g	ŋ
Taiof	∅	∅	k	g	ŋ
Hahon	∅	∅	.	.	n
Tinputz	∅	∅	∅	k	n
Teop	∅	∅	∅	g	n
Papapana	∅	∅	∅	g	ŋ
Banoni	g-; -g-, -∅-	g, ∅	g	.	ŋ
Piva	g-; -g-, -∅-	∅-; -g-, -∅-	.	g	ŋ
Uruava	∅	∅	∅	ŋ	ŋ-; -∅-, -ŋ-
Torau	∅	∅	∅, k	g	n
Mono-Alu	∅	∅	?	g	∅
Vaghua	g	g, ∅	k	g	g-; -ŋ-, -n-
Varisi	g, ∅	g, ∅	k	g	.
Ririo	g, ∅	∅	?	g	ŋ
Mbambatana	g, ∅	∅	.	g	ŋ
Sengga	∅	∅	.	g	ŋ
Lungga	g-, ∅-; -g-	∅	k	g	ŋ
Nduke	g, ∅	∅, g	.	g	ŋ
Roviana	g	∅	k	g	ŋ
Hoava	g	∅	k	g	ŋ
Vangunu	∅, h, g	∅	.	g	ŋ
Kia	g	g-; -∅-	k	g	ŋ; N/_i
Kokota	g, g; g/_C	g-; -∅-	.	g	ŋ
Laghu	g	g-; -∅-, -g-	k	g	ŋ
Blablanga	g, g	g	.	g	ŋ
Ghove	g; g/_C	g-, ∅-; -∅-	.	g	ŋ
Maringe	g; g/_C	g-, ∅-; -∅-	.	g	ŋ; N/_i

INTERPRETATION OF SYMBOLS

Banoni z = [dz].

Banoni b and g have merged and are interchangeable before rounded vowels (Lincoln 1976a:50).

The value of Uruava ŋ is unclear from the literature: it may be a mis-transcription of [g].

Voiced stops are prenasalised in Choiseul and New Georgia languages.

NOTES

1. This paper is a slightly revised version of a part of the paper presented at the Conference. I should like to thank two people who each contributed to the first version in very different ways. Darrell Tryon was a continual source of encouragement and a constructive critic. Jacques Guy introduced me to Simula, the programming language in which I wrote the programmes employed in the computer-aided analysis of sound correspondences upon which much of this paper is based. I also owe a methodological debt to the work at the University of Hawaii reported by Bender and Wang (1985), which indicates how the computer might be used as an aid in the analysis of sound correspondences. The present revision has benefited from the comments of Frank Lichtenberk, to whom I am also very grateful.
2. Although Bougainville is not part of the political entity of the Solomon Islands, it is geographically and culturally part of the Solomon Islands chain, as attested by its official name, the North Solomons Province of Papua New Guinea. The political boundary between Bougainville and the Shortland Islands was drawn only in 1899, when the boundary between the German and British protectorates was shifted. Hence this slightly curious name for an otherwise hard-to-name group of languages has geographical and a little historical justification.
3. Sources for the languages of the North-West Solomonic group were: for the languages of the Western Solomons Tryon and Hackman 1983, supplemented from Hackman n.d., from my fieldnotes on Mono-Alu, Mbambatana, Sengga and Roviana and from Waterhouse 1949 for Roviana, from Bosma 1980 for Maringe, and from Bosma 1981 for Kia and Maringe; for all Bougainville languages except Uruava and Piva, my own fieldnotes, supplemented by Todd 1978 for Nehan, by Allen and Baeso 1975 for Petats, by Allen and Allen 1965 and Allen 1971, 1978 for Hanahan, by Snyder and Snyder n.d. and Snyder 1981 for Teop, by Hostetler and Hostetler 1975 for Tinputz, by Lincoln 1976a for Banoni, by Rausch 1912 for Torau, and by Wheeler 1913a, 1913b, 1926 and Fagan 1979 for Mono-Alu. Data for Piva are from Lincoln 1976b and for the now extinct Uruava language from Rausch 1912.
4. I have adopted the convention that the abbreviations of the names of proto-languages begin not with P but with !. This is done both to distinguish proto-languages from languages whose names begin with P, and for reasons related to the use of the computer. Hence 'Proto-Oceanic' is abbreviated '!OC' rather than 'POC', 'Proto-North-West Solomonic' as '!NS' rather than as 'PNS'.
5. The following probably represents a further shared innovation: The third person plural disjunctive pronoun in languages of all North-West Solomons groups reflects a form !NS -ri[a], instead of one of the expected forms -di[a] or -ra. However, there is sufficient uncertainty about the form of Proto-Oceanic third person plural forms to warrant its exclusion for the moment.
6. In the examples, (proto-)morphemes which do not belong to the comparison are bracketed, whilst reduplications and phonological accretions are set off by a hyphen.
7. The system of glossing used in this paper is essentially the same as that used by Geraghty (1983:8-11).

8. Forms omitted from consideration because they appear not to be derived from the Proto-Oceanic forms reflected elsewhere in the North-West Solomonian region are:
- Haku a-ku *I* (possibly < !OC *ŋku *my*)
 Hahon, Teop e-an, Tinputz e-en *thou* (possibly metathesised forms reflecting !NS *e-r-(y)o)
 Mono-Alu maha *I*, maito *thou*, mani *we* (excl.), maan *you* (origin unknown)
9. Ysabel forms have lost *-m-; forms reflecting a possible !OC *kai *we* (excl.) are common in the Papua New Guinea region, and forms reflecting a possible !OC *kau *you* also occur in New Ireland. However, it is not relevant to the topic of this paper to consider whether these are independent parallel innovations or shared inherited forms.

REFERENCES

- ALLEN, Gerald [Jerry]
 1978 Halia verb morphology. M.A. dissertation, University of Texas at Arlington.
- ALLEN, Jerry
 1971 Tense/aspect and conjugations in Halia narratives. *Oceanic Linguistics* 10:63-77.
- ALLEN, Jerry and Conrad HURD
 1963 Unpublished survey word lists on file at SIL, Ukarumpa.
- ALLEN, Jerry and Janice ALLEN
 1965 *Halia language course*. Port Moresby: Department of Information and Extension Services.
- ALLEN, Jerry and Matthew BEASO
 1975 Petats phonemes and orthography. *Workpapers in Papua New Guinea Languages* 13:45-70.
- BENDER, Byron W. and Judith W. WANG
 1985 The status of Proto-Micronesian. In Andrew Pawley and Lois Carrington, eds *Austronesian linguistics at the 15th Pacific Science Congress*, 53-92. PL, C-88.
- BLUST, Robert A.
 1978 *The Proto-Oceanic palatals*. *Journal of the Polynesian Society* Monograph No.43. Auckland.
- BOSMA, David
 1980 Letasi cheke Maringe (An alphabet for Maringe). Solomon Islands Translation Advisory Group, Honiara. Mimeo.
 1981 Standardised alphabets for Santa Isabel languages. Solomon Islands Translation Advisory Group, Honiara. Mimeo.

FAGAN, Joel L.

- 1979 Preliminaries to a field study of Mono-Alu, Solomon Islands. M.A. dissertation, University of Hawaii.

GERAGHTY, Paul A.

- 1983 *The history of the Fijian languages*. Honolulu: University of Hawaii Press.

GREEN, R.C. and M. KELLY, eds

- 1972 *Studies in Oceanic culture history*, vol.3. Honolulu: Bernice P. Bishop Museum.

HACKMAN, B.D.

- n.d. Unpublished fieldnotes.

HALIM, Amran, Lois CARRINGTON and S.A. WURM, eds

- 1982 *Papers from the Third International Conference on Austronesian Linguistics*, vol.1: *Currents in Oceanic*. PL, C-74.

HOSTETLER, Roman and Carolyn HOSTETLER

- 1975 A tentative description of Tinputz phonology. *Workpapers in Papua New Guinea Languages* 13:5-44.

LEVY, Richard

- n.d. Languages of the southeast Solomon Islands and the reconstruction of Proto-Eastern Oceanic. University of Kentucky. Mimeo.

LINCOLN, Peter C.

- 1976a Describing Banoni, an Austronesian language of southeast Bougainville. Ph.D. dissertation, University of Hawaii.
1976b Banoni, Piva and Papuanization. PL, A-45:77-105.

MILKE, Wilhelm

- 1965 Comparative notes on the Austronesian languages of New Guinea. *Lingua* 14:330-348.

PAWLEY, Andrew

- 1972 On the internal relationships of eastern Oceanic languages. In R.C. Green and M. Kelly, eds *Studies in Oceanic culture history*, vol.3: 1-142. Honolulu: Bernice P. Bishop Museum.

RAUSCH, P.J.

- 1912 Die Sprachen von Südost-Bougainville, Deutsche Salomoninseln. *Anthropos* 7:105-134, 585-616, 964-994.

ROSS, Malcolm

- 1982 The development of the verb phrase in the Oceanic languages of the Bougainville region. In Halim, Carrington and Wurm, eds 1982:1-52.

SNYDER, David

- 1981 Teop dictionary. Second printing, level 7. Computer printout. Ukarumpa: SIL.

SNYDER, David and Ruth SNYDER

n.d. Teop orthographic considerations. Computer printout. Ukarumpa: SIL.

TODD, Evelyn M.

1978 A sketch of Nissan (Nehan) grammar. In Wurm and Carrington, eds 1978:1181-1239.

TRYON, D.T. and B.D. HACKMAN

1983 *Solomon Islands languages: an internal classification*. PL, C-72.

WATERHOUSE, J.H.L.

1949 *A Roviana and English dictionary*. Revised and enlarged by L.M. Jones. Sydney: Epworth Printing and Publishing House.

WHEELER, G.C.

1913a Nine texts in Mono speech. *Mitteilungen des Seminars für orientalische Sprachen* 16/1:1-48.

1913b A text in Mono speech. *Anthropos* 8:738-753.

1926 *Mono-Alu folklore (Bougainville Strait, Western Solomon Islands)*. London: George Routledge.

WURM, S.A., ed.

1976 *New Guinea area languages and language study*, vol.2. PL, C-39.

WURM, S.A. and Lois CARRINGTON, eds

1978 *Second International Conference on Austronesian Linguistics: proceedings*. PL, C-61.

ON DETERMINING THE EXTERNAL RELATIONSHIPS OF THE MICRONESIAN LANGUAGES

Frederick H. Jackson

0. INTRODUCTION

As it is used in this paper,¹ the term "Micronesian languages" refers to the indigenous languages of the Marshall and Gilbert Islands and to Kosraean (Kusaiean) and the Ponapeic and Trukic languages of the Caroline Islands. It thus includes Marshallese, Kiribati (Gilbertese), Kosraean, Ponapean, Mokilese, Pingilapese, Ngatikese, and the more than 30 dialects of the Trukic continuum, including, from east to west, Mortlockese, Lagoon Trukese, Puluwatese, Satawalese, Saipan Carolinian, Woleaian, Ulithian, Sonsorolese, Pulo Anna, Tobi, and the former language of the island of Mapia, about 100 miles north of Irian Jaya. Located within geographical Micronesia but excluded from the linguistic grouping are Yapese, the Polynesian Outliers Kapingamarangi and Nukuoro, and the non-Oceanic Austronesian languages Palauan and Chamorro.

In my recent dissertation (Jackson 1983), I present evidence that the Micronesian languages, as just defined, constitute a distinct subgroup of Oceanic. The purpose of the present paper is to review proposals that have previously been made regarding the external relationships of Micronesian languages in the light of what is now known about those languages, and also to present some additional data which may be pertinent to the discussion.

The first section of this paper briefly summarises some aspects of Proto-Micronesian (PMC) which have been reconstructed. Section 2 demonstrates the likelihood that Nauruan forms a higher order subgroup with Micronesian, a subgroup which I have named Greater Micronesian. Section 3 summarises and evaluates evidence which has been presented in support of the linkage of Micronesian with the languages of Northern Vanuatu (Grace 1955, 1964, 1971; Pawley 1972, 1977), San Cristobal and Malaita (Blust 1984), and the Admiralty Islands (Smythe 1970), and concludes that none of the claims is supported sufficiently. The concluding section briefly presents some additional data which may or may not be relevant to the determination of the genetic relationships of Micronesian and suggests that the uncovering of conclusive evidence must probably await fuller understanding of both the other Oceanic languages and the ancestral Proto-Oceanic language.

1. PROTO-MICRONESIAN

Table 1 lists the Micronesian reflexes of Proto-Oceanic (POC) consonant phonemes.² These reflexes, together with lexical data not reported here,

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 201-238. *Pacific Linguistics*, C-94, 1986.

© Frederick H. Jackson

Table 1: Reflexes of POC consonant phonemes among Micronesian languages³

POC	*p	*mp	*ŋp	*k	*ŋk	*q	*m	*ŋm	*n	*ñ	*ŋ		
PMC	*f	∅ ¹	*p	*pw ²	*k	*x	∅	*m	*m	*m ²	*n	*ñ	*ŋ
KSR	∅	∅	p	f	k	k, ∅ ³	∅	m	w, m ⁴	n	∅	∅	ng
PCMC	*f	∅	*p	*pw	*k	*x, ∅	∅	*m	*mw	*n	*ñ	*ñ	*ŋ
KIR	∅	∅	b ⁵	b, b ¹⁶	k, ∅ ⁷	∅	∅	m	m, m ¹⁶	n	n	n	ng
PWMC	*f	∅	*p	*pw	*k	*x, ∅	∅	*m	*mw	*n	*ñ	*ñ	*ŋ
MRS	∅	∅	p	b	k, kw ⁸	∅	∅	m	m ⁸ , m ⁸	n, n ⁸	n, n ⁸	n, n ⁸	ñ
PTK-PP	*f	∅	*p	*pw	*k	*x, ∅	∅	*m	*mw	*n	*ñ	*ñ	*ŋ
PON	p, ∅ ⁹	∅	p	pw	k	r, ∅ ¹⁰	∅	m	m ¹⁰	n	∅	∅	ng
MOK	p, ∅ ⁹	∅	p	pw	k	r, ∅ ¹⁰	∅	m	m ¹⁰	n	∅, n ¹¹	∅, n ¹¹	ng
PTK	*f	∅	*p	*pw	*k	∅	∅	*m	*mw	*n	*ñ	*ñ	*ŋ
TRK	f	∅	p	pw	k, ∅ ¹²	∅	∅	m	m ¹²	n	n	n	ng, n ¹³
MRT	f	∅	p	pw	k, ∅ ¹⁴	∅	∅	m	m ¹⁴	n	n	n	ng
PUL	f	∅	p	pw	k, ∅ ¹⁴	∅	∅	m	m ¹⁴	n	n	n	ng
STW	f	∅	p	pw	k, ∅ ¹⁴	∅	∅	m	m ¹⁵	n, l ¹⁵	n, l ¹⁵	n, l ¹⁵	ng
CRL	f	∅	p	pw	gh, kk, ∅ ¹⁶	∅	∅	m	m ¹⁶	l	l	l	ng
WOL	f	∅	p	b, pw ¹⁷	g, kk ¹⁸	∅	∅	m	m ¹⁷	l, n ¹⁹	l, n ¹⁹	l, n ¹⁹	ng
PUA	d ²⁰	∅	p	pw	k, kk ¹⁸	∅	∅	m	m ¹⁸	n	n	n	ng
SNS	f	∅	p	pw	k, kk ¹⁸	∅	∅	m	m ¹⁸	r	r	r	ng
ULI	f	∅	p	b	g, kk ¹⁸	∅	∅	m	m ¹⁸	l	l	l	ng
MAP	f, v	∅	p	b, ∅	g, k	∅	∅	m	n, m ²¹	n	∅	∅	n

¹POC *p was lost in PMC before round vowels.
²POC *mp and *m became labiovelars in PMC before round vowels and in some other forms.
³PMC *x is usually lost in KSR before *u.
⁴PMC *mw is reflected as KSR m before a vowel that is lost word finally.
⁵KIR b is a lax voiceless stop.
⁶The symbol b' is not written before round vowels.
⁷PMC *k is sometimes lost in KIR morphemes that also reflect PMC *t.
⁸MRS variants are predictable from following vowels.
⁹PMC *f is often lost in PP before *i.
¹⁰PMC *x is r before a and lost elsewhere.
¹¹PMC *ñ is MOK n after high vowels.
¹²PMC *k is often reflected as TRK s before *i, and regularly as ∅ before low and mid vowels when not also preceded by a high vowel.
¹³PMC *ŋ normally becomes TRK n before *i.
¹⁴PMC *k is normally lost before non-high vowels when not also preceded by a high vowel.
¹⁵STW l and n are apparently in free variation.
¹⁶CRL gh is a velar fricative; PMC *k is normally lost before non-high vowels when not also preceded by a high vowel.
¹⁷WOL pw represents a geminate stop.
¹⁸PMC *k is a velar fricative singly and kk when geminate.
¹⁹WOL n represents a geminate nasal.
²⁰PUA d is a dental fricative.
²¹MAP ñ is a velar nasal; MAP m occurs in final position.

Table 1 (cont'd)

POC	*l	*d	*R	*nd	*nt	*t	*s (PEO) *nj	*ns	*j	*w	*y
PMC	*l	*r	∅ ²²	*c		*t	*d	*z		*w	∅ *T ²³
KSR	l	l	∅	sr ²⁴		t, s ²⁵	t, s ²⁵	∅	∅	∅	s
PCMC	*l	*r	∅	*c		*t	*d ²⁶	*d ²⁶	*w	∅	*T ²³
KIR	n	∅	∅	r		t, ∅ ²⁷	r	r	w	∅	t
PWMC	*l	*r	∅	*c		*t	*d	*d	*w	∅	*T ²³
MRS	l, } ⁸	r	∅	d ²⁸		j ²⁹	t	t	w	∅	j ²⁹
PTK-PP	*l	*r	∅	*c		*t, ∅ ³⁰	*d	*d	*w	∅	*T
PON	l	r	∅	t ³¹		s, ∅ ³²	d ³³	d ³³	w, u ³⁴	∅	s
MOK	l	r	∅	s		j, ∅ ^{32, 29}	d ³³	d ³³	w, u ³⁴	∅	j ²⁹
PTK	*l	*r	∅	*c		*t, ∅ ³⁰	*d	*d	*w	∅	*T
TRK	n	r	∅	ch ³¹		s, ∅ ³²	t	t	w	∅	s
MRT	l	r	∅	sh, ²⁴ ch ³⁵		s, ∅ ³²	t	t	w	∅	s
PUL	l	r	∅	r, ³⁶ ch ³⁵		h, s, ∅ ³⁷	t	t	w	∅	h
STW	l, n ¹⁵	r	∅	rh, ³⁶ ch ³⁵		s, ∅ ³²	t	t	w	∅	s
CRL	l	r	∅	sch, ²⁴ tch ³⁵		s, ∅ ³²	t	t	w	∅	s
WOL	l, n ¹⁹	r, ch ³⁵	∅	sh, ²⁴ ch ³⁵		t, s, ∅ ³⁸	t	t	w	∅	s, t ³⁹
PUA	n	l	∅	s		t, d, ²⁰ ∅ ³⁸	t	t	w	∅	d, ²⁰ t ³⁹
SNS	r, n ¹⁹	l	∅	s		t, d, ⁴⁰ ∅ ³⁸	t	t	w	∅	d, ⁴⁰ t ³⁹
ULI	l	r	∅	c ³¹		t, s, ∅ ³⁸	d ⁴¹	d ⁴¹	w	∅	s, t ³⁹
MAP	l	r	∅	c, t		t, t ^v , j, h, ∅ ⁴²	t	t	v(?)	(?)	(?)

²²The conditions for loss of *R are not clear

²³It is not certain that *T was distinct from *t before PTK-PP. Reconstructed only before *i, *u, *e.

²⁴This phoneme is a retroflex spirant.

²⁵KSR reflects PMC *t and *d as s before *i and *e.

²⁶PCMC merged PMC *d and *z as PCMC *d.

²⁷PMC *t is sometimes lost in KIR morphemes which also reflect PMC *k (see fn.7).

²⁸MRS d is a palatalised retroflex trill.

²⁹MRS, MOK j are palatalised dental stops.

³⁰PMC *t was apparently lost in a very few forms in PTK-PP before non-low vowels.

³¹This phoneme is a retroflex stop.

³²PMC *t is generally lost before *u, *i, and *e, and reflected as s (MOK j) elsewhere. Several exceptions.

³³PON, MOK d are alveolar stops.

³⁴Orthographic u represents /w/ when not followed by a vowel.

³⁵This symbol represents a geminate retroflex stop.

³⁶These symbols represent retroflex approximants.

³⁷PMC *t is lost before *u, *i, and *e, otherwise reflected as h singly and s when geminate. Several exceptions.

³⁸PMC *t is usually reflected as t before *a and s (PUA, SNS d) elsewhere. Many exceptions, especially in WOL.

³⁹PTK-PP *T is reflected as t before *u.

⁴⁰SNS d is a dental fricative.

⁴¹ULI d is a dental fricative.

⁴²PMC *t is reflected as MAP t before word-final *a which is deleted and as MAP t^v or j before *a elsewhere; it is lost before other historical word-final vowels and reflected as MAP h before non-low vowels which are retained.

indicate that the following 11 phonological innovations are shared by all Micronesian languages. To my knowledge, the same cannot be said of any other language.

- (1) Split of POC *mp into PMC *p and *pw, with PMC *pw occurring before round vowels regularly and before non-round vowels in certain forms.⁴
- (2) Split of POC *m into PMC *m and *mw, with a pattern of reflexes similar to that for POC *mp (see (1) above).⁵
- (3) Loss of POC *p before round vowels.
- (4) Split of POC *R into PMC *r and Ø under unidentified conditions.⁶
- (5) Merger of POC *nt and *nd as PMC *c, which was almost certainly a retroflex obstruent.
- (6) Merger of PEO *z (POC *ns) and POC *j as PMC *z.⁷
- (7) Merger of PEO *s (POC *s) and PEO *nj⁸ as PMC *d.
- (8) Reasonably consistent reflexes of POC *n as PMC *ŋ in the environment /*a__*i in the following five lexical items: PMC *kaŋi *sharp* (< PEO *kani); PMC *tapaŋi *help, assist, support* (< PEO *tampani); PMC *maŋi *think, remember, recall* (< PEO *mani); PMC *ta(ŋ,n)i *from, source* (< POC *tani); PMC *ka(ŋ,n)i *eat* (vt) (< POC *kani). This change is not attested in any MC reflexes of the following etyma: PMC *raani *day* (< POC *daqani); PMC *tani *skin disease* (< PEO *tani); PMC *pani *sea cucumber* (< PEO *pani); and PMC *wa(a)ni *pumice*.
- (9) Spirantisation of POC *t before *i.
- (10) Loss of POC *q.
- (11) Loss of POC *y.⁹

Detailed reconstruction of the PMC grammatical system has only just begun. However, it nonetheless seems possible at this time to make a brief statement about basic sentence structure and also to reconstruct a few closed grammatical systems for PMC.

1.1 PMC sentence structure

It appears highly probable that PMC transitive sentences were strictly SVO, but that speakers had the option in intransitive sentences of postposing the subject noun phrase. In that event, however, the subject pronouns, which were syntactically a part of the verb phrase, remained in preverbal position. It also appears very probable that PMC speakers made extensive use of processes of focus and topicalisation whereby the relevant noun phrase was moved to the front of the sentence, typically leaving a pronominal trace. In all sentences which included a specific verb phrase, aspect markers, negative morphemes, and most aspectual adverbs probably were positioned between the subject pronoun and the verb.

Similarly to many other Oceanic languages, however, PMC also included sentences consisting only of two noun phrases, where the second phrase served to identify or make a statement about the first. The following hypothetical reconstruction provides an example of this kind of sentence.

- (1) *mwaane na aramata-ni dakau
man that person-of reef-island
That man (is) of a reef island.

In all PMC sentences, demonstrative morphemes almost certainly followed the head noun, but it is not yet clear whether numerical constructions also followed the noun or preceded it. Harrison (n.d.) has demonstrated the likelihood of a PMC article *te, also reflected in Nauruan (see Section 2), which preceded the noun in some constructions.

1.2 PMC personal pronouns

The probable PMC system of personal pronouns is shown in Table 2. It is perhaps noteworthy that only singular and plural forms are reconstructed. Although evidence for apparently dual and trial forms is found in Ponapeic, Marshallese, and Kosraean, their structures differ, and at present it seems more probable that those forms represent separate developments of those languages.

Reconstruction of the first person plural inclusive and exclusive and second person plural focus and object pronouns is somewhat problematic, as it appears to be necessary to reconstruct doublets for the six forms. In the case of the first person plural inclusive forms (PMC *ki(t,c)a) the difference reflected is only in the grade of the medial consonant, with Marshallese and Kosraean reflecting oral grade POC *t and the other languages reflecting nasal grade *nt. Similar variation in grade in this etymon is also reflected elsewhere in Oceanic (e.g. in Fijian, where Bauan keda reflects the nasal grade while Lauan keta reflects oral grade).

The reconstructed doublets for the PMC second person plural focus and object pronouns (*kamii, *kamwu '2pl. focus and object pronouns') and for the first person plural exclusive forms (*kamami, *kami 'lpl.exc. focus and object pronouns') would appear to represent a more serious problem. A first step towards a solution is provided by Harrison's persuasive argument that in PMC there were no distinct plural object pronouns and that the plural focus pronouns were able to function as pronominal objects in the object noun phrase position (Harrison 1978). Harrison's proposal in effect restricts the problem of the doublets to only the focus pronoun set. A second step toward a solution to the problem is suggested when it is observed that one member of each of the focus pronoun doublets is identical with the corresponding subject pronoun reconstructions: cf. PMC *kami 'lpl.exc. focus, object, and subject pronoun', *kamwu '2pl. focus, object, and subject pronoun'. The replacement of earlier subject pronouns by focus pronouns has occurred to various extents in all MC languages except Kiribati and the Trukic languages, with the two meanings under discussion apparently showing the greatest extent of replacement. It appears likely that this replacement process may reflect a confusion between plural subject and focus pronouns that occurred in the proto-language, leading to historically subject pronouns being reinterpreted as alternate focus pronoun forms. If so, then the pre-Micronesian focus pronouns in these meanings were presumably *kamami lpl.exc.' and *kamii '2pl.', and the doublet forms *kami 'lpl.exc.' and *kamwu '2pl.', which need to be reconstructed for PMC as both focus and subject pronouns, functioned in pre-Micronesian only as subject pronouns.¹⁰

Table 2: Micronesian personal pronouns¹

	PMC	KSR	KIR	MRS	PTK-PP	PON	MOK	PTK ²
FOCUS PRONOUNS								
1sg.	*ŋau	nga	ngngai	ñā	*ŋau	ngehi	ngoahi	*ŋaŋu
2sg.	*koe	(kom) ³	ngkoe	kwe	*koe-ñā	kowe	koawoa	*koe-ñā
3sg.	*ia	(el) ⁴	ngaia	e	*ia	ih	ih	*ia
1pl.inc.	*ki(t,c)a	kuht	ngaira	kōj	*kica	kit-	kihs	*kica
1pl.exc.	*kamami	(kitacl) ⁵	-	kōmmem ⁶	*kamami	(kiht) ⁸	-	*kaamami
	*kami	-	-	kōm ⁷	*kami	-	kam- ⁹	*kami
2pl.	*kamii	-	ngkamii	komi ⁷	*kamii	-	-	*kamii
	*kamwu	kom-tacl ¹⁰	-	kor ⁶	*kamwu	kumw-	kanw- ⁹	-
3pl.	*ira	el-tahl ¹⁰	ngaiia	er	*ira	ihr	ihr	*ira
SUBJECT PRONOUNS								
1sg.	*ú	-ll	i	i-	*ú	i	-ll	*ú
2sg.	*ko	-ll	ko	ko-	*ko	ke	-ll	*ko
3sg.	*e	-ll	e	e-	*e	e	-ll	*e
1pl.inc.	*t(i,e)	-ll	ti	je-	*T(i,e)	-ll	-ll	*Ti
1pl.exc.	*kami	-ll	-	-ll	*kami	(se) ¹²	-ll	*kami ¹³
2pl.	*kamwu	-ll	kam'	-ll	*kamwu	-ll	-ll	*kamwu ¹⁴
3pl.	*ra	-ll	a	re-	*re	re	-ll	*re
OBJECT PRONOUNS								
1sg.	*-ai	-yuc	-ai	-eō	*-ai	-ye	-15	*-ai
2sg.	*-ko	-15	-ko	-eok	*-ko	-yuk	-15	*-ko
3sg. (anaph.)	*-a	-15	-a	-	*-a	-ø	-15	*-a
1pl.inc.	*ki(t,c)a	-15	-(i)ra	-15	*kica	-15	-15	*-kica
1pl.exc.	*kamami	-15	-	-15	*kamami	-15	-15	*-kamami
	*kami	-15	-	-15	*kami	-15	-15	*-kami ¹⁶
2pl.	*kamii	-15	-15	-15	*kamii	-15	-15	*-kamii
3pl. forms:								
anaph.hum.	*ira	-15	-ia	-15	*ira	-15	-15	*-ira
anaph.inan.	*-ni(?) ¹⁷	-	-	-	*-ni(?)	-	-	*-nini
3 (before noun phrases)	*-xi	-	-i	-(?)	*-i	-	-	*-i
POSSESSIVE PRONOUNS								
1sg.	*-xu	-k	-u	-i	*-j	-j	-i	*-j
2sg.	*-mwu	-m	-m'	-ŋ	*-mwu	-mw	-mw	*-mwu
3sg.	*-ña	-ø	-na	-n	*-ña	-ø	-ø,-n ¹⁸	*-ña
1pl.inc.	*-ca	-sr	-ra	-d	*-ca	-t-	-s-	*-ca
1pl.exc.	*-mi	(-k-tacl) ¹⁹	-	-m	*-mi	(-t) ²⁰	-m	*-mi ²¹
	*-mami ²²	-	-	-	*-mami ²²	-	-	*-mami
2pl.	*-mii	(-m-tacl) ²³	-mii	-mi	*-mii	(-mw-) ²⁴	(-mw-) ²⁴	*-mii
3pl.	*-(i)ra	-l-tacl ²⁵	-ia	-er	*-ira	-Vr-	-Vr-	*-ira

¹Standard orthography of the sources is used throughout. PON, MOK h indicates the preceding vowel is long.

²All Trukic languages reflect the PTK reconstructions except as noted.

³KSR kom is a reflex of the PMC 2pl. form.

⁴KSR el is a reflex of the PMC 3pl. form.

⁵KSR kitacl is a reflex of the PMC 1pl.inc. form with a trial suffix.

⁶Western dialect form in MRS.

⁷Eastern dialect form in MRS.

⁸PON kiht is a reflex of the PMC 1pl.inc. form.

⁹These MOK forms are roots for the dual and trial series.

¹⁰These KSR forms also reflect earlier trial suffixes.

¹¹The focus pronoun is used for this function.

¹²PON se is a reflex of the PMC 1pl.inc. form.

¹³The medial consonant of PTK *kami is only reflected in MAP. All other TK languages reflect the type *kai.

¹⁴The medial consonant of PTK *kamwu is only reflected in MAP. All other TK languages reflect the type *kau.

¹⁵The focus pronoun is used for this function.

¹⁶PTK *kami is only reflected in this function in TRK and a few closely related TK dialects.

¹⁷Tentatively reconstructed for PMC and PTK-PP on the basis of TK and external evidence.

¹⁸Regular MOK reflexes which are dependent on the stem-final vowel (see Table 1).

¹⁹KSR -k-tacl reflects the PMC 1sg. form with an earlier trial suffix.

²⁰PON -t reflects the PMC 1pl.inc. form.

²¹PTK *-mi is only reflected in TRK and a few closely related TK dialects.

²²Reconstructed for PMC and PTK-PP on the basis of TK and external evidence.

²³KSR -m-tacl reflects the PMC 2sg. form with an earlier trial suffix.

²⁴Reflexes of the PMC 2sg. form.

²⁵KSR -l-tacl reflects the PMC 3pl. form with an earlier trial suffix.

Also quite problematic is the determination of how third person objects may have been marked on the verb in PMC. Harrison (1978) presents a strong case for there having been two transitive suffixes in PMC: a morpheme *-a, not the personal pronoun, which was used with expressed singular noun phrase objects, and the more expected morpheme *-i, which was used with pronoun objects, plural noun phrase objects (including plural focus pronouns), and also anaphorically for plural non-human objects. As part of his argument, Harrison demonstrates the great probability that the third person plural pronoun *ira referred only to humans. My own research on the Trukic languages of Micronesia, while generally confirming Harrison's reconstruction of the two transitive suffixes and his proposed restriction on the meaning of *ira, also indicates that, at the Proto-Trukic level at least, there were four distinct forms for marking third person objects. Specifically, there were the expected pronoun objects *-a and *-ira, with the former probably used anaphorically for singular objects and the latter used anaphorically for plural human objects, but there was also a form *-nini, which was used anaphorically for plural inanimate objects, and a form *-i, distinct from the transitive suffix, which was used before both singular and plural expressed noun phrase objects.

There is some evidence that the PTK object suffix *-i is reflected elsewhere in Micronesia in Kiribati and Marshallese (Harrison 1978:1077-1078), but there is no other evidence in MC of the PTK *-nini suffix. In spite of this lack of attestation, however, it appears very possible that both Trukic suffixes should be reconstructed for PMC on the grounds that they represent continuations of earlier forms. Geraghty (1983:158-159) has pointed out that a number of South-east Solomonian languages, including Kwara'ae,¹¹ appear to reflect a type *-(ŋ)ki 'plural inanimate suffix on verbs' which, in the nasal grade, is a likely cognate of PTK *-i and the KIR and MRS forms in Micronesia, suggesting PMC *-xi. The case for a PMC source of PTK *-nini is considerably more tentative, but can still be presented. Churchward (1940:17) reports a Rotuman form ne as a "plural indefinite article" which is never used "elsewhere than after a verb", and Ivens (1933, 1937) apparently considered Ngela to have a postverbal "anticipatory object" form ni which was "used of things" (Ivens 1933:153). (See Jackson 1983: 244-245 for discussion.) Both of these forms would be regular reflexes of a type *ni, and, if cognate, may suggest a source for PTK *-nini.

1.3 PMC postverbal directional morphemes

Seven postverbal directional enclitics are reconstructed for PMC, as follows:

	PMC	PTK	PON	MRS	KIR	KSR
<i>thither, toward hearer</i>	*watu	*wa(t)u	wei	waj	wati	-
<i>outwards, out to sea</i>	*otu	*wo(t)u	iei	-	-	wot
<i>downwards, downwind, west</i>	*zio	*diwo	di	to	rio	i
<i>upwards, upwind, east</i>	*zake	*dake	da	tak	rake	yak
<i>hither, toward speaker</i>	*mai	- ¹²	-	-	mai	ma
<i>away from speaker</i>	*lako	*lako	la	ɔk	nako	lac
<i>inwards, inland, ashore</i>	*loŋo	*loŋo	long	lōñ	-	-

PMC *otu reflects the type *potu, which is attested in Kove potu *outside, beyond the reef, seaward side of island*, and Kove otu *go out, leave an island (by canoe), appear* (Chowning 1973 and personal communication); in the Rotuman postverbal hofu

toward the coast (from inland) (Churchward 1940); and in Marau Sound wou *away at sea* (cf. Marau Sound wau *out at sea* < PEO *watu). It is also clear that putative PMC *loŋo is cognate with, e.g. Fijian yaloga *inside*, Bugotu i-longa *landwards*, Vaturanga longa *ashore, inland, south*, Kove loŋa *enter, come ashore, interior of New Britain*, and the Rotuman postverbal loŋa *toward the interior (of an island)*.¹³ Since neither Kiribati nor Kosraean reflects the form, however, it is not clear whether the innovative final vowel attested in the other Micronesian languages developed in PMC or at some later date.

All other PMC directionals are obvious reflexes of POC forms.

1.4 PMC demonstrative morphemes

All Micronesian languages have extremely complex demonstrative systems, and the correspondences among them are not completely clear. However, it appears possible at this time to reconstruct the following three demonstrative roots for PMC: *e 'near speaker'; *na 'near addressee'; *oe 'away from speaker and addressee; known referent'. The first of these morphemes is apparently cognate with Kadavu Fijian and Tongan ē *here*, near speaker (Geraghty, personal communication), and the second is widely spread throughout Oceanic. The only likely cognate that I have been able to locate for *oe, however, is Nauruan -oo, in the same meaning, a fact which provides some support for grouping Nauruan with Micronesian (see Section 2).

1.5 Possessive classifiers

All Micronesian languages except Kiribati have relatively extensive sets of possessive classifiers which are used in the possession of alienable nouns. Strikingly, however, only four possessive classifiers can be reconstructed with any confidence for PMC, and it is not certain that even those forms functioned in the proto-language in the same way as possessive classifiers in the modern languages. The reconstructed forms are:

	PMC	PTK	PON	MRS	KIR	KSR
<i>general object</i>	*a-	*aa-	a(a)-	a(a)-	a-	- ¹⁴
<i>drinkable object</i>	*nima-	*nima-	nima-	(n,l)ime-	-	nihmac
<i>edible object</i>	*na-	- ¹⁵	- ¹⁵	- ¹⁶	-	na-
<i>offspring</i>	*natu-	*natu-	nei-	naji-	-	nahtuh

PMC *a- has cognates in Nauruan, Polynesian (PPN *qa-), Ambrym, Sesake, and several languages of the south-east Solomons, including Faghani, Oroha, and Arosi. PMC *natu also has widespread cognates, but not, to my knowledge, ones which function as possessive classifiers. PMC *na- *edible object*, while reflected within Micronesian only in Kosraean, apparently has external cognates with similar meaning in Fijian (Nabukelevu, Kadavu: Geraghty, personal communication) and Santa Cruz (Codrington 1885), and is reconstructed for PMC on that basis. PMC *nima- *drinkable object* is formally identical with a PMC (and Nauruan) verb meaning *to drink*, but otherwise appears unique to Micronesia.

Harrison (1981) has recently suggested that the modern Kiribati grammar may reflect the proto-language possessive system more closely than do the grammars of any of the other modern languages. He reconstructs PMC *a as what he terms a

'possessive article', and suggests that other possessive classifiers in the modern Micronesian languages derive from a grammatical process of nominal apposition in the proto-language. Such a process occurs in modern Kiribati, as shown by the following phrase:

- (2) nima-u te ran
drink-my ART water
my water (to drink)

Harrison's arguments are too detailed to enter into here. They are very persuasive, however, and although I have differed from Harrison in reconstructing, primarily on the basis of external evidence, four PMC possessive classifiers, it is not absolutely clear that this decision is the correct one. Moreover, even if it is correct, it is very likely that the other modern possessive classifiers evolved in the way suggested by Harrison.

1.6 PMC numbers and counting classifiers

Harrison and Jackson (1984) have shown that Micronesian number roots from *two* through *nine* reflect the well known POC reconstructions, with the exception that PMC *fa(a)- *four* fails to reflect the final syllable of POC *pati.¹⁷ Two distinct forms for *one* are reconstructed for PMC, however. In serial (abstract) counting, the PMC root for *one* was probably a reflex of POC *(n)sa, but in specifying an amount of countable objects a unit prefix PMC *te- *one* was used together with the appropriate counting classifier. In a similar fashion, the other number roots were also prefixed to an appropriate counting classifier to indicate a specific number of objects.

Although several modern Micronesian languages attest an extremely large number of such classifiers - for example, both Kiribati and Trukese have more than 100 - only five may be securely reconstructed for the proto-language. These are: PMC *-ua *general objects* (< POC *pua *fruit*); PMC *-manu *animate objects* (< POC *manu(k) *bird, creature*); PMC *-cau *thin flat objects* (< POC *(n)dau *leaf*);¹⁸ PMC *-ḡaulu *units of tens* (< POC *ḡa-pulu *ten*); and PMC *-pwukua *units of hundreds*. As noted, all but the last of these forms reflect well known POC etyma. The last, identical in form with PMC *pwukua *knee, node, joint*, may be cognate with the type ^kp^wu^yu- *knee* attested in the Banks Islands and Northern Vanuatu (Tryon 1976) and with Fijian buku *humped, knotted*, PPN *puku *protuberance, lump, swelling* (and cf. Roviana puku-a *to tie, knot*). It appears very likely that counting in the Proto-Micronesian culture in part involved the enumeration of knots in a rope.

In addition, one other counting classifier was probably present in PMC. This somewhat tentative reconstruction is PMC *-ḡaratu *units of thousands*, a reflex of POC *Ratu(s) *hundred*, which is reflected in Micronesia in PTK *-ḡaratu and, with problematic loss of the final syllable, in KIR ngaa, both with the meaning *unit of thousands*.

Harrison and Jackson (1984) also reconstruct additional classifiers for PMC: *k(u,i)di, *lopwa, *dep(u,i), and *nena, each with the meaning *units of high powers of ten*. Under the subgrouping relationships assumed in this paper, however, it is no longer possible to claim that these forms were present in PMC.

1.7 PMC interrogative morphemes

Strictly speaking, interrogative forms are not grammatical morphemes. They do form a relatively closed set, however, and are thus of value in comparative reconstruction.

The PMC interrogatives are reconstructed in Table 3. Although all of them are also attested outside of Micronesia, the complete system is still of interest. Among the reconstructed forms, PMC *fi(d,z)a *how much, how many* clearly reflects POC *pija, while PMC *(mee-)zaa *what?* almost certainly derives from a compound consisting of a cognate of PPN *meqa *thing* and a reflex of PEO *za(a) *what?* (Geraghty 1983). The same compound, but with a different meaning, is reflected in Fijian meca *thing*. Sources for the other PMC forms are somewhat less obvious.

Table 3: Interrogative forms in Proto-Micronesian

	PMC	KSR	KIR	MRS	PON	MOK	PTK
<i>how much, how many?</i>	*fi(d,z)a	(ekahsr) ¹	ira	-	depe ²	doapoa ²	*fida-
<i>what?</i>	*(mee-)zaa	meac	raa	ta	da ³	da ³	*mee-daa
<i>who?</i>	*tau	suc	antai	(won)	(ihs)	inje	*i-tau
<i>when?</i>	*ŋaiza	ngac	ni-ngaira	nāāt	-	ngehd	*i-ŋaeda
	*ñaiza	-	-	-	i-ahd	-	*i-ñaeda
<i>which?</i>	*-faa	-	(raa)	-	(i)a	(i)a	*-faa
<i>where, how, which place?</i>	*i-faa	-	iaa	(ewi)	-	-	*i-faa
<i>where?</i>	*ia(a)	-yac	iaa	ia	ia	ia	*i-iaa

¹Non-cognate forms are provided in parentheses.
²These forms apparently reflect an occurrence of consonant metathesis.
³The Ponapeic language Ngatikese attests a form mahda *what?*, which reflects the reconstructed compound.

The somewhat unexpected *t which is reconstructed in the PMC form for *who?* (cf. POC *(n)sai *who?*) is apparently also attested in Gedaged itai (Dempwolff n.d.:30-31) and in Nambel and Morouas ita (Tryon 1976). A palatal nasal in a form for *when?* is clearly reflected in several Santa Ysabel languages in the Solomon Islands, including Bugotu (Tryon and Hackman 1983), as well as in Trukic and Ponapean. PMC *-faa *which?*, *i-faa *where, how, which place?* appear to be cognate with PPN *fe(a,e) *where?* (Biggs 1979), reflexes for which include Niuean fee *where, which, when?*, Rennellese hea *which, what, where?*, and Tongan 'ifē *which?*, fēfē *in what way, of what sort, how?* (cf. also Faghani iafee, Vaturanga iava *where?* (Codrington 1885), both of which suggest still wider distribution for the form). PMC *ia(a) *where (to)?* is probably cognate with Seimat iia and Wuvulu ia *where?* in the Admiralty Islands (Smythe 1970; Blust, personal communication) and with the form ia which Lynch and Tryon (1985) report for several Southern Vanuatu languages, although the authors assign it to their putative Proto-Central Oceanic *pia.¹⁹

1.8 Conclusion

Although, as noted above, the grammatical reconstruction of Proto-Micronesian is only just beginning, we now have sufficient knowledge of it to permit tentative

comparisons to be made with other languages and language groups, and to evaluate other such comparisons which have previously been proposed.

2. THE RELATIONSHIP OF NAURUAN TO THE MICRONESIAN LANGUAGES: A TENTATIVE CASE FOR A GREATER MICRONESIAN GROUPING

We turn now to an examination of the quite meagre data available on the Nauruan language in the hope of clarifying its relationship to the Micronesian languages. Both Bender (1971) and Nathan (1973) expressed the tentative hope that Nauruan would prove to be aligned with Micronesian, but were unable to identify any specific shared innovations. To my knowledge, no linguistic work has been done on Nauruan since Nathan in the early 1970s,²⁰ but our increased knowledge of Micronesian languages and of PMC makes it more possible now to interpret the Nauruan data that exist.

Nathan (1973:483) suggests that the Nauruan (NAU) consonant inventory is as follows:

p	pw	t	k	kw	: voiceless stops
b	bw	d	g	gw	: voiced stops
m	mw	n	ŋ	(ŋw)	: nasal stops
		r, ʔ	j	w	: approximants

Nathan believes that the proposed distinction between voiceless labial stops (/p,pw/) and voiced ones (/b,bw/) may be one of length on the underlying level, but he is not certain (Nathan 1973, n.d.a). The other proposed voicing contrasts are apparently more clearly phonemic (Nathan n.d.a). The phoneme /r/ has flapped and trilled allophones, with the flap typically occurring before stressed vowels (Nathan n.d.a). The other liquid phoneme /ʔ/ is "trilled, slightly long and accompanied by a buzzing sound [Nathan] can approximate by partial devoicing in medial position. Initially it sounds a little like [dr]" (Nathan n.d.a:6). Both /j/ and /w/ are close glides.

The data presented in Nathan 1973 and in his unpublished grammatical sketch (Nathan n.d.a) and word list (Nathan n.d.b) indicate that the probable inherited NAU reflexes of the POC consonant phonemes are as follows (with PMC reflexes also included for reference):²¹

Table 4: Nauruan reflexes of POC consonants (with PMC reference)

POC:	*p	*mp	*ŋp	*t	*nt	*d	*nd	*q
PMC:	∅,*f	*p, pw	*pw	*t	*c	*r	*c	*∅
NAU:	∅,p	b,p,bw	bw	∅	t	r	ʔ	∅
POC:	*k	*ŋk	*s	*ns	*j	*l	*R	
PMC:	*k	*x	*d	*z	*z	*l	∅,*r	
NAU:	∅,k,kw,g	k,∅	d,t	d,t	?	∅,r	∅,r	
POC:	*m	*ŋm	*n	*ñ	*ŋ	*w	*y	
PMC:	*m,*mw	*mw	*n	*ñ	*ŋ	*w	∅	
NAU:	m,mw	mw	n	n	ŋ,mw	w,∅	∅	

In discussing the correspondences set forth in Table 4, I shall first establish that NAU is not a member of the Micronesian subgroup as it was defined in Section 1 above. I will then list evidence which suggests that NAU is a coordinate member of a larger grouping, which will be termed Greater Micronesian.

2.1 Evidence that Nauruan is not a member of the Micronesian subgroup

As the correspondences in Table 4 show, NAU fails to reflect one of the phonological innovations which together define the Micronesian subgroup: the merger of POC *nt and *nd as PMC *c. NAU clearly reflects nasal grade POC *nt as t in kotá 'lpl.inc. object pronoun' and -tá 'lpl.inc. possessive pronoun' (while losing oral grade *t), but reflects nasal grade *nd as ʃ in, e.g., ʃAMWA- *forehead* (< POC *(n)da(ŋ)ma), ʃaa- *blood* (< POC *(n)daRa), ʃen *fresh water* (< POC *(n)danu). Oral grade POC *d is clearly reflected as NAU r in, e.g., -ra '3pl. possessive pronoun', roo *two*, marama- *moon*, oor *lobster*. The fact that NAU retains the historical distinction between *nd and *nt while all Micronesian languages merge them is clear grounds for not including NAU within the Micronesian subgroup.

Additional support for this conclusion is also apparent:

- (1) NAU appears to attest an r reflex of POC *R in aar *current* (< POC *qaRus(a)), while all MC languages show loss of POC *R in that etymon. In all other known comparisons, however, PMC and NAU agree in showing loss or retention of POC *R (see below).
- (2) Forms reconstructed for PMC with *f have cognates in NAU where both ∅ and p correspond with PMC *f: NAU e-pee *stone*, PMC *fatu; NAU e-een *woman*, PMC *fa(i)fine; NAU ji *select, choose*, PMC *fili. It appears likely that NAU p reflects nasal grade POC *mp and that oral grade POC *p is lost in NAU, and if so, the NAU and PMC reflexes of POC *patu *stone* do not agree in grade.
- (3) NAU also fails to reflect two important PMC lexical innovations: the replacement of POC *R by a velar nasal in PMC *maiŋi *left side* (cf. POC *mauRi > NAU -mow), and the replacement of final *i by *a in PMC *wakara *root* (cf. POC *wakaR(i) > NAU awori-).

2.2 Evidence for a Greater Micronesian subgroup

Although it is clear that NAU is not a member of the Micronesian subgroup, there is a rather considerable body of data which suggests that NAU and the Micronesian languages are closely related to each other. This evidence is provided here:

- (1) Like the MC languages, NAU has developed labiovelar reflexes of both POC *mp and *m before round vowels. Strikingly, of the 18 comparisons between PMC and NAU which I have identified where a labiovelar is reconstructed for PMC, NAU also attests a labiovelar in 17. The only exception is NAU mmar *string of beads*, which appears to correspond with PMC *mware *garland, lei*.
- (2) PMC and NAU agree in showing retention or loss of POC *R in nine of the ten comparisons that have been identified (the exception is noted above). Merger of POC *R and *d is attested in PMC *wakara, NAU awori- *root* < POC *wakaR(i), and both PMC and NAU attest loss of *R in the following comparisons: PMC *caa, NAU

*aa- *blood* < POC *(n)daRa; PMC *kuita, NAU *giigoo octopus* < POC *kuRita; PMC *ñañewa, NAU *nene yesterday* < POC *ñoRa; PMC *maa, NAU *maa ashamed* < POC *maRa; PMC *dou *spear, stab, inject*, NAU *tou cut, stab* < POC *(n)saRu; PMC *cui, NAU *e-ʔ bone* < POC *nduRi (Blust 1978); PMC *pau *wing, arm*, NAU *e-be wing* < POC *(m)paRu *wing* (Blust, personal communication); PMC *pwauzu, NAU *bwoodi- nose* < earlier *pwa-Runsu (see Section 3.2 for discussion of the last reconstruction).

(3) NAU has cognates of PMC *-ua 'counting classifier for general objects', *-manu 'counting classifier for animate and human objects', and *-cau 'counting classifier for thin flat objects' in the same meaning and function.

(4) NAU a- 'general possessive classifier' is cognate with PMC *a- in the same meaning and function.

(5) NAU *nim drink, to drink* and *nima- 'possessive classifier for drinkable objects'* are cognate with PMC *nima and *nima- in the same meanings and functions, and are apparently unattested elsewhere.

(6) NAU -oo 'demonstrative root: away from speaker and hearer' is almost certainly cognate with PMC *oe in the same meaning and function, but is also apparently not attested elsewhere.

(7) NAU *ji where?* is very probably cognate with PMC *ia(a) *where (to)?*, other cognates of which are very rare in Oceanic.

(8) NAU *eʌ:kʌŋ sharp* appears to be cognate with PMC *kaŋi *sharp*, which represents an innovation from PEO *kani.

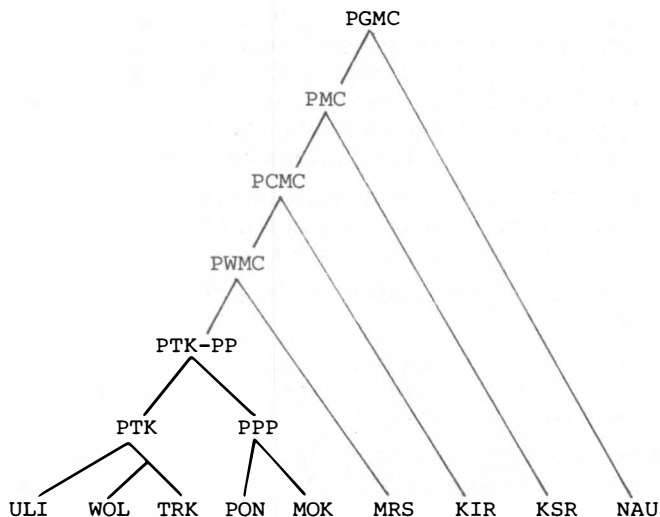
(9) NAU *e-ʔ bone* is almost certainly cognate with PMC *cui, reflecting an earlier *nduRi which is otherwise only attested in the Admiralties and some non-Oceanic Austronesian languages (Blust 1978).

(10) NAU *bwiʔibwiʔ white* is cognate with the otherwise apparently innovative PMC *pwece(pwece) *coral lime, white*.

(11) NAU *mwi tooth* may reflect the type *ŋii *tooth* reconstructed in Jackson (1983:388) only for Proto-Central Micronesian because of the absence of a cognate form in Kosraean. NAU *mw* appears to reflect an earlier velar nasal in at least one other comparison: NAU *bwumw night* < POC *ŋpoŋi *night*. If the NAU form is cognate with PCMC *ŋii, presumably that form must then be reconstructed for PMC as well. Since POC *(n,ñ) *upon tooth* is securely reconstructed, this NAU and MC etymon would appear to represent an innovation.

(12) NAU *bwoodi- nose* is cognate with PMC *pwauzu in the same meaning. Blust (1984) shows that this etymon is also attested in several Cristobal-Malaitan languages, and suggests that it represents a shared innovation of his proposed Cristobal Malaitan-Micronesian subgroup (see Section 3.2 for discussion of Blust's subgrouping claim). Even if Blust is correct, however, the NAU form provides evidence for linking NAU with Micronesia.²²

From such a small corpus of data as is available on NAU, the number of arguments that have been identified for subgrouping NAU with Micronesian is quite impressive, albeit not conclusive. Pending considerably more data on NAU, however, I shall tentatively propose such a subgroup, which I will provisionally term Greater Micronesian. Although the reconstruction of Proto-Greater Micronesian must wait until another time, the following genetic tree will describe the relationship which I am proposing between NAU and the Micronesian languages.



3. EARLIER PROPOSALS REGARDING THE EXTERNAL RELATIONSHIPS OF THE MICRONESIAN LANGUAGES

This section reviews and discusses proposals which have previously been published regarding the possible external relationships of the Micronesian languages and demonstrates that the evidence is conclusive for none of them. The first such proposal to be considered is the suggestion by Grace (1955, 1964, 1971) and Pawley (1972, 1977) that the Micronesian languages may be linked with languages of Northern Vanuatu. The second is Blust's suggestion that the Micronesian languages subgroup with languages of San Cristobal and Malaita in the south-east Solomon Islands (Blust 1984), and the third is Smythe's suggestion that there are linguistic similarities between the languages of the Admiralty Islands and those of Micronesia (Smythe 1970).

3.1 The Grace-Pawley hypothesis

In 1955, George Grace published what he termed "tentative conclusions" about the subgroups of "Eastern Malayo-Polynesian", now more commonly called Oceanic (Grace 1955). Among those conclusions, Grace stated that "the membership of this grouping [Micronesian, but including Yapese] in the New Hebrides-Banks subgroup is highly probable, but not certain". Other proposed members of the New Hebrides-Banks subgroup were Polynesian, Fijian, Rotuman, and all the languages of the New Hebrides and Banks and Torres Islands.

Although he has never published the evidence on which he based this conclusion about Micronesian, Grace does not appear to have changed his opinion. In 1964 he writes:

With the exception of Palauan and Chamorro, I believe that all of the languages of Micronesia belong to a single group which has its closest relations with languages of the New Hebrides, although not with those languages which are most

closely related to Rotuman, Fijian, and Polynesian. The languages of Yap and Nauru are generally regarded as quite aberrant, but I believe, nevertheless, that they belong to this Micronesian grouping. If I am right, these relations indicate a movement from the New Hebrides to Micronesian [sic].

(Grace 1964:367)

Somewhat later, Grace (1971) states that he is "rather less certain that Yapese fits into the Micronesian group". From the foregoing, it is clear that Grace's Micronesian group is now equivalent to the Greater Micronesian group proposed in this paper.²³

The only Micronesian language that Andrew Pawley considered in his investigation of 31 Eastern Oceanic languages was Gilbertese (Kiribati) (Pawley 1972). In exploring the possible subgrouping relationships of that language, Pawley states (1972:134) that "there appear to be no grounds for treating Gilbertese as a Southeast Solomonian language", but suggests that there is some evidence for including it in North Hebridean-Central Pacific (HC), a group which he proposes as including Fijian, Polynesian, and languages of the Banks, Torres, and northern New Hebrides Islands.²⁴ This proposed evidence will be discussed below.

Pawley (1977) redefines the term "Eastern Oceanic" to refer strictly to his earlier North Hebridean-Central Pacific group. He remains uncertain whether Micronesian languages should be included in the redefined EO group, but states, "There appears to be a distinct chance that they can be" Pawley asserts that "a language can be placed in a subgroup by virtue of its possession of one or more of a list of diagnostic features" (Pawley 1977:4 [my emphasis]). He then proposes a list of ten such features for inclusion in EO. I will list and discuss these proposed features in relation to the Micronesian languages.

(1) Loss of POC *R. Pawley observes that this criterion is "problematical" in that "*R is not lost in all regions" of the redefined EO, but merges with POC *d in "one small region of North Hebridean-Central Pacific, namely the Banks Islands languages" (Pawley 1977:8). Therefore, he concludes that *R was lost in "a dialect of the proto-language" (1977:9).

Pawley's assumption that only the Banks Islands in the New Hebrides have non-zero reflexes of *R has been proven untenable by both Tryon (1976) and Geraghty (1978). Indeed, as I shall demonstrate below, it now appears that most, if not all languages of the Northern New Hebrides have non-zero reflexes of *R in at least some forms. Moreover, even if Pawley's criterion were diagnostic of all the proposed EO languages, Micronesian would not meet that criterion. Although Pawley refers without discussion to Milke's (1958) claim that "*R is lost in all the languages which would now be termed Nuclear Micronesian and in Yapese" (Pawley 1977:8), appearing to imply that he agrees with this claim, he was explicitly aware in 1972 of the fact that Dyen (1949:425) had observed that some instances of *R are reflected in at least Trukese as r (Pawley 1972:134-135).²⁵ We now know that the merger in some lexical forms of *R and *d is attested in all MC languages except Kiribati, which has lost both *R and *d.

Although it is clear that Pawley's criterion is not strictly met by the Micronesian languages (or by many of the languages in the New Hebrides), another somewhat more complex possibility suggests itself. Since both Micronesian and New Hebridean languages attest loss of *R in some lexical items but merger with *d in other items, if it could be shown that Micronesian languages attest loss of *R in the same lexical items where loss is attested in a New Hebridean language

or group of languages and that the lexical items which reflect merger of *R and *d are also the same in both groups of languages, it would be striking evidence for subgrouping them together.

In an effort to identify a language or language group in the New Hebrides which has the same pattern of reflexes of *R as the Micronesian languages, the data in Tryon 1976 have been searched for reflexes of reconstructions with *R that are also attested in PMC. The results of this search are shown below, with an indication for each New Hebridean language whether it is spoken in the Banks Islands (BA), on Maewo (MA), on Pentecost (PE), on Santo (SA), on Malekula (ML), on Efate (EF), or in the Shepherds (SH).

I. Comparison where all NH languages and PMC agree in attesting loss of *R

- (A) POC *Ruŋma *house*: PMC *(u,i)mwa; Mota (BA) imwa, Raga (PE) imwa, Nambel (SA) ima, Uripiv (ML) na-im, Nguna (EF) na-sumwa.

II. Comparisons where all NH languages and PMC agree in showing merger of *R and *d

- (B) POC *qapaRa *shoulder*: PMC *afara; Dixon Reef (ML) mbot-βera-ni *shoulder*, Benour (ML) ni-βara- *hand, wing*, Bongabonga (SH) na-mbarau- (?) *shoulder*.
 (C) POC *maRaqaŋ *light in weight*: PMC *maraara; Mosina (BA) mamara, Marino (MA) marara, Valpei (SA) mararaha, Aulua (ML) memar.
 (D) POC *wakaR(i) *root*: PMC *wakara; Mota (BA) ɣariw, Nambel (SA) ar-na, Dixon Reef (ML) wari, Peterara (MA) xori.

III. Comparisons where PMC loses *R and NH languages exhibit mixed reflexes

- (E) POC *kuRita *octopus*: PMC *kuita; Mota (BA) wirita, Mosina (BA) wirit, Raga (PE) xuita, Malo (SA) xwita, Uripiv (ML) na-it, Sesake wi:ta, Nguna (EF) wita. (With only Banks Islands languages attesting a non-zero reflex of *R.)
 (F) POC *ñoRa *yesterday*: PMC *ñañewa; Mota (BA) ananora, Mosina (BA) le-nor, Merlav (BA) nano, Seke (PE) nano, Nambel (SA) nanof (?), Rerep (ML) nenof, Sesake nanoβa (?), Nguna (EF) nanoβa (?). (With only some Banks Islands attesting a non-zero reflex of *R.)²⁶
 (G) POC *paRi *stingray*: PMC *fai; Mota (BA) βar, Mosina (BA) βer, Raga (PE) fari, Tolomako (SA) βari, Narang (SA) afai, Uripiv (ML) -βi, Sesake βai, Nguna (EF) βai. (With non-zero reflexes of *R occurring in the Banks Islands, Pentecost, and part of Santo.)
 (H) POC *(n)daRa *blood*: PMC *caa; Mota (BA) nara.²⁷ Mosina (BA) noro, Raga (PE) ndaxa-, Nambel (SA) ndrae-, Uripiv (ML) ndra-, Sesake -nda-, Nguna (EF) -da-.²⁸ (With only Banks Islands languages attesting a non-zero reflex of *R.)
 (I) POC *(m)paRu *wing*: PMC *pau *arm, wing*; Toga (Torres) pərpəriŋə *wing*, Raga (PE) xapau-, Wusi (SA) apau-, Axamb (ML) xapö-, Port Vato (Ambrym) a^mbau-, Sesake na-βaru-, Nguna (EF) na-liβaru- *wing*. (If this is a valid comparison, the correspondences are very puzzling. They appear to show Sesake and Nguna retaining a non-zero reflex of *R which is otherwise only attested in the Torres Islands. It is possible that Nguna, Sesake na-pau *shoulder* is the regular reflex of the POC form (cf. Uri (ML) ^mbiri^mbari-, Rano (ML) ^mbar^mbari-, Dixon Reef (ML) ^mbot-βerani, Port Vato (Ambrym) ^mbar-βelo-, all with the meaning *shoulder*). If so, two NH reconstructions may be necessary: one a reflex of POC *(m)paRu *wing* and the other a hypothetical *(m)par(i,u) *shoulder*. In addition, an assumption must be made that Nguna and Sesake switched

the meaning of the two forms so that, e.g., Sesake na-βaru- *wing* actually reflects the hypothetical *(m)par(i,u) *shoulder*, and Sesake na-pau *shoulder* reflects POC *(m)paRu *wing*. The only other apparent solution is to assume an early semantic split in reflexes of POC *(m)paRu into separate lexical items for *wing* and *shoulder*, with *R being separately and distinctly lost or merged with *d by the individual languages.)

- (J) POC *Ruqa *neck*: PMC *ua; Raga (PE) mwaxoro (?), Peterara (MA) xaua, Nokuku (SA) alo-, Axamb (ML) -xalua-. This comparison is also problematic, as Tryon (1976) does not recognise l as a reflex of *R or *d in any of these languages. However, given the fact that Nokuku is recognised by Tryon as losing *k in some (undetermined) environments, the NH comparisons strongly suggest the reconstruction of a form *kaCua *neck*, where C represents a liquid which was lost in Peterara. The gloss and the POC reconstruction suggest almost as strongly that the liquid was *R. If so, then there is also a good possibility that the final disyllable of the Raga form represents the putative *ka(-)Rua. If this comparison is viable, however, then it represents a case of *R being lost only in Maewo, being merged with *d in Raga (PE), and merged with *l in languages of Santo and Malekula.

IV. Comparisons where PMC merges *R with *d and NH languages exhibit mixed reflexes.

- (K) POC *takuRu *back*: PMC *takuru; Wetamut (BA) tawuru-, Wusi (SA) ta'u, Marino (MA) tayu, Mae (ML) taxu. (With only Banks Islands languages attesting a non-zero reflex of *R.)
- (L) POC *meRa *red, reddish*: PMC *mera; Mota (BA) memea, Mosina (BA) ya-meme, Raga (PE) memea, Marino (MA) memea, Tutuba (SA) memea, Aulua (ML) miel (?), Malua Bay (ML) i-mel (?), Lelepa (EF) mila (?), Sesake miala (?). It is possible that the Aulua, Malua Bay, Lelepa, and Sesake forms in this comparison in fact reflect an earlier *meRa-la, with the last syllable an unidentified morpheme. If that is the case, this comparison represents an instance of POC *R being lost throughout the NH, but merged with *d in MC (cf. within MC, MRS mir *red-coloured, of reddish coconuts or sky*, PON mer *rusty, corroded*, MOK mehr *stained from coconut husk or fruit*, KIR mea *rust, grey, reddish yellow colour*).

Table 5 (overleaf) summarises the MC and NH reflexes of POC *R in the 12 lexical items.

Table 5: Reflexes of POC *R in 12 lexical items for MC and NH languages

POC	PMC	Banks, Torres	Maewo	Pente- cost	Santo	Male- kula	Shepherds	Efate
*Ruŋma	∅	∅	-	∅	∅	∅	-	∅
*qapaRa	r	-	-	-	-	r	r	-
*maqaRa	r	r	r	-	r	r	-	-
*wakaR(i)	r	r	r	-	r	r	-	-
*kuRita	∅	r	∅	∅	∅	∅	-	∅
*ñoRa	∅	r,∅	∅	∅	∅	∅	-	∅
*paRi	∅	r	r	r	r,∅	∅	-	∅
*(n)daRa	∅	r	-	∅(?)	∅	∅	-	∅
*(m)paRu	∅	r(?)	-	∅	∅	∅	-	r(?)
*Ruqa	∅	-	∅	r	l	l	-	-
*takuRu	r	r	∅	∅	∅	∅	-	-
*meRa	r	∅	∅	∅	∅	∅(?)	-	∅(?)

In the first four comparisons summarised in Table 5, there is agreement among PMC and all the NH languages reflecting the forms on whether *R was lost or merged with *d. This fact may be supportive of the suggestion that PMC subgroups with languages of the New Hebrides, but it does not help in determining a specific dialect group in the New Hebrides from which PMC might have derived. However, when the remaining eight comparison sets are examined to attempt to identify such a group, the evidence is extremely contradictory. This is true even if the problematic comparisons involving POC *(m)paRu and *Ruqa are discarded.

The only way in which reflexes of POC *R can be used as evidence for subgrouping PMC with a New Hebridean dialect is by assuming that the pattern of reflexes that are now attested for *R in those languages are the same as the reflexes which occurred at the time of the putative separation of PMC from the New Hebrides. That is, for PMC and a New Hebridean dialect to be subgrouped on the basis of reflexes of *R in specific lexical items, it must be assumed that *R was either lost or merged with *d in specific lexical items in the proto-language of PMC and that dialect. If *R existed as a separate phoneme in the ancestral language of PMC (i.e. was not lost or merged with *d until later), then the lexical pattern of its reflexes is irrelevant to subgrouping. When the data summarised in Table 5 are examined, however, it is clear that even if PMC did derive from a dialect in the New Hebrides, *R must have been a distinct phoneme at the time of separation. The reflexes of POC *meRa *red* are sufficient to demonstrate that fact, as, under the scenario that we are for the moment assuming, *R must have been lost in the NH languages after the ancestor of PMC had separated, and *R must have been distinct in PMC at that time for it to later merge with *d.

When the reflexes in the remaining cognate sets shown in Table 5 are also examined, the same conclusion is reached: that if the source for PMC was a dialect area of the New Hebrides, *R must have still been a distinct phoneme when the separation occurred. If *R was not distinct at that time, and had already been lost or merged with *d in individual lexical items, then the evidence of the *takuRu cognate set indicates that PMC must have derived from the Banks Islands, which is the only place besides Micronesia where *R merged with *d. On the other hand, the evidence of the cognate sets for *kuRita, *paRi, and *(n)daRa indicates that PMC, which has lost *R in reflexes of those forms, could not have derived from a Banks Island language because *R is merged with *d there.

In the context of his review of Tryon 1976, Geraghty (1978) makes the following observation:

... the disappearance of *R is gradual, increasing from north to south within the New Hebrides, with no apparent phonological conditioning. Nor is the irregularity confined to the New Hebrides: the liquid *R seeps away in sporadic losses from one end of Oceania to the other. Already on the threshold of Oceania, PAN *maRi *hither* and *kaRi *speak* unexpectedly become POC *mai and *kai *say*; and there are many POC doublets suggesting occasional loss or change of *R in daughter languages. In all of the New Hebrides, *R is lost in the forms *Rumwa *house*, *meRa *red*, and (probably) *viRa *kind of taro*. Between Vanua Lava and Santa Maria in the Banks Islands, *R disappears from *noRa [sic] *yesterday*, and *kaRati *bite* follows suit before we reach Merlav. Between Merlav, the last of the Banks Islands, and Maewo, *kuRita *squid* loses its *R and half way down Maewo the same happens to *suRi *bone*. South of Pentecost Island, *R is lost from *vaRi *stingray*, and the loss is shared also in all of Santo (except for two languages in the North West). Finally, a central area comprising West Ambrym, South West Epi, and most of the Shepherd Islands leads the field in loss of *R, deleting it also in *mauRi *left hand*; and somewhere between the New Hebrides and Fiji we must draw another line, for all of the New Hebrides languages retain *R in *tuRi *sew* and *kaRu *swim*, while all Fijian and Polynesian languages, and Rotuman, lose *R in all instances.

(Geraghty 1978:81-82)

Although another explanation has been proposed for the variable reflexes of *R,²⁹ I believe that the most likely explanation is that the changes affecting *R have been gradually diffusing through the lexicons of Oceanic languages in the manner described by Wang (1969, 1979) for other languages and characterised by him as "lexical diffusion" (see also Jackson 1983, 1984 for a discussion of lexical diffusion in Micronesia). It is possible, as Geraghty (1978) implies, that loss of *R began to diffuse through the lexicon as early as in POC. In any event, it eventually affected all eligible lexical items in the Central Pacific languages and apparently, as Pawley (1977:8) states, in New Caledonia, the Loyalty Islands, and the Admiralty Islands as well. Perhaps simultaneously, the competing change of merger of *R with *d has also been gradually spreading through the eligible lexical items.³⁰ Ross (1977) shows that it has occurred in many (but not all) languages of the Sepik and Western Madang coast in New Guinea, and Bradshaw (1978) and Ross (1979) indicate that it probably also took place in other New Guinea languages. Indeed, as Geraghty (1978:85n) points out, although the south-east Solomons languages merge *R with *l normally, Sa'a ahuri *conch* (for expected **ahuli < POC *tapuRi) suggests an irregular merger of *R and *d in that form, as does the Proto-Southeast Solomonic reconstruction *bara *fence* < POC *(m)paRa.

As we have seen, both of these diffusing changes are attested in MC and NH languages. While it is clear that one change or the other had affected all the eligible lexical items by the time of the dispersal of PMC, the occurrence of Nauruan aar *current* and the PMC reconstruction *auda (< POC *qaRus(a)) suggests that at least one lexical item was unchanged in Proto-Greater Micronesian. Similarly, *R must have been a distinct phoneme in NH languages in at least the majority of lexical items at the time when the ancestor of the Micronesian

languages separated from the New Hebrides, if indeed it did so. As stated earlier, the fact that all MC and NH languages have the same reflexes of *R in *Ruŋma *house*, *qapaRa *shoulder*, *maRaqaŋ *light in weight*, and *wakaR(i) *root* may indicate that the diffusing changes occurred in a proto-language ancestral to both language groups, but it is also possible that this agreement is entirely coincidental. (For a discussion of far more numerous instances of coincidental agreement of a lexically diffused change, see Jackson 1984.) Thus, although the MC and NH reflexes of *R do not contradict the possibility that the two language groups subgroup together, they also do not provide tangible support for that claim.

(2) Replacement of POC *kami '1pl.exc. focus pronoun' by PEO *kamami. Pawley's second criterion for inclusion within the PEO group would appear to be met by MC languages. Although the Eastern dialect of Marshallese, Trukese, and Mokilese appear to attest the type *kami in this meaning, all other Trukic languages and also the Western Marshallese dialect clearly reflect *kamami. It is questionable, however, how valid this feature is for purposes of subgrouping. As Blust (personal communication) has pointed out, Palauan kemam '1pl.exc. emphatic pronoun' also reflects the type *kamami, and the same type is also attested in other non-Oceanic Austronesian languages. It is very likely, therefore, that the fact that many of the proposed EO languages attest this form is indicative of a retention of an earlier form, rather than a shared innovation.

(3) Replacement of POC *ka- 'food or drink possession marker' by *ma- 'drinkable possession' and *ka- 'edible possession'. As shown in Section 1.5, PMC almost certainly had the possessive classifier *nima- 'drinkable possession' and *na- 'edible possession'. The former was related to the PMC verb *nima *to drink*, while the latter reflects an etymon also reflected in non-EO languages.

(4) Replacement of POC *pai *where?* by PEO *p(i,e)a. It was observed in Section 1.7 that PMC *-faa *which?* (cf. PMC *i-faa *where, how, which place?*) is probably cognate with PPN *fe(a,e) *where?*, which would suggest that MC languages meet this criterion. However, it was also noted that both Faghani and Vaturanga of the south-east Solomons also appear to attest the same etymon, and neither of those languages is included in Pawley's proposed EO grouping. Moreover, Ivens (1937: 1090) records Florida *ivia where, what, how?* in addition to Florida *ivei where, how, which, what?*, with both of Pawley's POC and PEO forms appearing to be reflected. Thus, it appears very probable that PEO *p(i,e)a was not an innovation of the putative EO group.

(5) "Recutting of POC *(Ca)kini 'prepositional verb marking instrumental-causative relation' as *ki-ni. This yielded a new preposition *ki; the *-ni segment was reanalysed as part of the following object person-marker. Apparently *ki also took on a dative-marking function in PHC." There appear to have been two PMC reflexes of POC *(Ca)kini: PMC *-aki 'agentless passive suffix on verbs' and PMC *-(a)kin- 'remote transitive suffix'. Neither of these reflects the proposed PEO innovations.

(6) Reflection of PEO *teqe 'preverbal negative particle'. Kā'eo (n.d.) has proposed that the following MC negative morphemes reflect a PEO *taqe: MRS ja-, PON sa-, KSR suh-, se-, TRK se, WOL te, tai, MOK ja-, PUL há, PUA taai. In fact, though, WOL te 'prohibitive negative' reflects a PTK-PP *de, and the other forms apparently reflect PMC *ta(i) 'preverbal negative'. The forms of the MRS, PON, and MOK morphemes strongly indicate a short vowel rather than the long vowel

which would be predicted from Kā'eo's reconstruction, and the consonant reflexes of the Trukic and Ponapeic languages point to an initial *ta-. The fact that KIR attests a form tai 'negative imperative' provides further support for this analysis (Harrison, personal communication). The proposed comparison of PEO *teqe with PMC *ta(i) has too many irregularities to be considered as subgrouping evidence.

(7) Reflection of PEO *tika(i) 'negative verb'. Pawley (1972) suggests that KIR tiaki *no* might reflect this reconstruction by metathesis, but, as Blust (1984) has pointed out, it appears that the KIR form is bimorphemic: KIR ti *only*, aki *not*. There are no other potential reflexes in MC.

(8) Reflection of PEO *ma 'preverbal particle marking past or non-future tense'. Although the system by which aspect was marked in PMC has thus far eluded reconstruction, there is no evidence in any MC language of a reflex of *ma in this meaning.

(9) Reflection of PEO *(η)kai 'preverbal conjunctive particle relating clauses in temporal succession'. Kā'eo (n.d.) proposes that MRS nē *when, if*, KSR nge *and finally*, TRK, WOL, ULI, CRL nge *and, but* reflect this form in Micronesia. However, no MC language reflects either *k or *ŋk as a velar nasal. Thus, this comparison must also be discarded.

(10) Reflection of PEO *(n)tewa *one*. There is also no evidence in MC of a reflex of this reconstruction, although Kā'eo (n.d.) has suggested that the following forms reflect it: MRS juo-n, KSR se, TRK eew, WOL se-, PUL ye-, CRL et. se-, and PUA deei, dawa. In fact, the MRS and TRK forms reflect the number-general classifier compound PMC *te-ua, the unit prefix of which is also reflected by the cited KSR, WOL, and PUL forms, and by CRL se-. CRL et (correctly eet) reflects the serial counting form for *one*, PMC *-za (< POC *(n)sa), with the prefix *e-. The forms which Kā'eo cites for PUA are something of a mystery, as Oda's grammar and lexicon of the language provide only the unit prefix de- (< PMC *te-) and the compound de-ow *one (general object)* (< PMC *te-ua) (Oda 1977), and Kā'eo does not provide her source.

In sum, the comparison of data from the MC languages with Pawley's suggested criteria for inclusion in the putative Eastern Oceanic grouping provides no basis for including MC within that group. Although the Micronesian and New Hebridean reflexes of *R do not deny such a grouping, they also do not support it. Similarly, the fact that MC and EO languages both attest reflexes of PEO *kamami and *p(i,e)a loses its subgrouping value when it is recognised that there are witnessesses of both forms which are external to EO. The other seven proposed criteria are not met.

3.2 Blust's Malaita-Micronesian hypothesis

Blust (1984) proposes that Micronesian languages are genetically related to those of San Cristobal and Malaita in the South-east Solomon Islands. This proposal has several merits, but, I submit, is again not conclusive.

Because Blust is unable to identify any "clear phonological or syntactic innovations shared exclusively by the Nuclear Micronesian and Cristobal-Malaitan languages ..." (Blust 1984:101-102), the evidence which he presents in support

of his hypothesis consists of lexical comparisons. He presents 28 such comparisons, of which the first two are the most impressive. Twenty-three of the comparisons are claimed to reflect Proto-Malaita-Micronesian (PMMC) innovations, a claim which Blust strengthens by stating that he searched through the published lexical material on more than 200 other Oceanic languages and was unable to find cognates. In this section, I will present and discuss eight of Blust's proposed comparisons, including all of the ones which he considers to provide the strongest evidence for the proposed grouping.

The Cristobal-Malaitan (CM) languages which are cited by Blust are 'Are'are (AA), Arosi (ARS), Bauro (BRO), Kwaió (KWO), Lau (LAU), Sa'a (SAA), and Ulawa (ULW).

(1) Proto-Malaita-Micronesian (PMMC) *pwaRusu *nose*

Reflected in AA, ARS, KWO, LAU, SAA, and in the firm PMC reconstruction *pwauzu *nose*, this is a most impressive comparison. (I would suggest altering the reconstruction to *pwaRusu, though, on the grounds that PMC *z reflects earlier *ns, while the CM languages are indeterminate between *s and *ns.) If it were the case that no languages external to the proposed grouping attested the form, it would provide strong evidence for Blust's subgrouping proposal. There appear to be some possible external cognates which Blust has missed, however. For example, Yapese p'eethnguu-n *his nose* almost certainly reflects the same etymon. Although there is considerable evidence of borrowing between Yapese and the neighbouring MC language of Ulithi, the Ulithian cognate is bōōdē- *nose*, with vowels and an initial consonant that are quite distinct from their counterparts in the Yapese word. Moreover, although no Yapese reflexes of reconstructed forms can as yet be termed "regular", Bradshaw (1975) lists several instances where the Yapese dental fricative th is a reflex of earlier *(n)s, and also provides other information that would suggest that Yapese p'eeth- is a directly inherited reflex of an earlier *pwaRusu. He states that "one source for [Yapese] glottalised consonants is the coalescence of a glottal stop and some other consonant". The Yapese forms ruŋ'-ag *hear* (< PAN *DeŋeR) and pi' *give* (< PAN *beRey) indicate PAN *R is sometimes reflected as Yapese glottal stop. The replacement of *R by glottal stop in *pwaRusu would provide the source for the glottalised Yapese p'. The presence of the syllable -nguu-, which is certainly not of Ulithian provenance, on the Yapese form also suggests that the Yapese form is not borrowed.

Although the Yapese form is the only likely cognate of *pwaRusu external to the proposed MMC of which I am aware, there are some data in Tryon 1976 which suggest earlier sources for distinct morphemes *pwa- and *Ru(n)su, both with meanings related to *nose*. For example, Sowa of Pentecost has a form bwa-ŋsu *nose* which appears to reflect earlier *pwa- compounded with a reflex of POC *(ŋ)isu *nose*. In addition, several Maewo, Santo, and Malekula languages may reflect an earlier *R(i,u)(n)su in their forms for nose: cf. Marino, Peterara līsu-, Malmariv, Lametin nalsu-, Morouas nasu-, Lorediakarkar, Shark Bay lusu-, Tangoa, Araki ŋalusu-, Dixon Reef ngarsi. Although Tryon does not recognise l reflexes of *R in any of these languages, we have seen in our discussion of New Hebridean reflexes of POC *Ruqa *neck* that it is possible that such reflexes may occur. Certainly the forms cited for Morouas and Dixon Reef in the above comparison suggest a proto-segment distinct from *l.

If my interpretation of these data is correct, and *pwa- and *Ru(n)su were separate morphemes present in a language ancestral to the languages of the New Hebrides and of Cristobal-Malaita, perhaps even POC, it becomes more possible that the compound *pwaRusu could have developed independently in MC and CM.

Alternatively, as the Yapese form indicates may have been the case, the compound may have developed at a much earlier date and simply been retained in CM and MC.

(2) PMMC *masawa (better: *mansawa) *sea beyond the reef; deep sea*

This reconstruction is reflected in AA, ARS, BRO, KWO, LAU, SAA, and in PMC *mazawa. If it is an innovation, then it would also be powerful evidence for the relationship adduced by Blust. Again, however, it seems likely that it is a continuation of an earlier form. Kove, a language which Blust does not appear to have searched, is reported by Chowning (1973:232) as attesting the form maroani *deep water*, which appears compatible with the proposed PMMC reconstruction: cf. Kove *tari sea* < POC *tansi(k), *waro sun* < POC *qanso, and *Nera when?* < POC *ŋaqija for other instances of Kove *r* < POC *ns. Although the final syllable of the Kove form may appear problematic, Chowning (personal communication) states that /i/ is the expected vowel following POC final consonants which are retained in Kove, so it is possible that the *n* in Kove maroani in fact reflects such a final consonant. POC final consonants are regularly lost in this environment in both MC and CM languages. In partial support of the analysis of the Kove form presented here, Chowning (personal communication) has also brought the Kove form maroani-havu *high seas* to my attention, where Kove havu means *middle*.

Blust himself notes Yapese mathaaw *deep sea, open ocean*, but observes correctly that Jensen (1977) considers it to be a borrowing from Ulithian madaw (Blust 1984:110). It may be worth noting, however, that the Yapese correspondences would not be irregular for a directly inherited form. Blust also suggests that Roviana *masa beach, seashore*, Manam *marā deep blue sea*, and Bugotu *masa* [sic] *deep, of sea; the deep sea* may be related, and indicate an earlier *masa, without the final syllable of the putative PMMC form. This observation needs modification, however, as Manam *r* regularly reflects POC *ns (cf. Manam *ore paddle* < POC *ponse, and *raoa son- or daughter-in-law; parent-in-law* < POC *nsawa spouse (Lichtenberk 1977)), and because Ivens (1940) cites no form *masa* for Bugotu. Instead, Ivens reports Bugotu *maha to be deep of sea; the deep sea* and *tahi maha open sea*, both of which also point to *ns. Thus, both the Manam and Bugotu forms strongly suggest earlier *mansa *deep sea*. Indeed, since Levy (n.d.) indicates that Bugotu regularly loses POC *w, it is not inconceivable that at least the Bugotu forms reflect *mansawa.³¹

In light of the above, it can not be certain that the PMMC reconstruction represents an innovation.

(3) PMMC *maŋo *breath, to breathe; fontanel*

The primary problem with this proposed comparison is that only one of the six CM reflexes which Blust suggests attests the gloss *fontanel* - LAU *maŋo-na pulse, beat of heart; fontanel; lungs; life, soul, spirit; wind, breath* - while on internal evidence the only gloss that can be reconstructed for PMC *maŋo is *forehead, fontanel*; cf. KIR *mango fontanel*, MRS *mōŋ pate of head, fontanel*, KSR *mahngo top, head, ridge, crown (of the head)* (3ps), MOK *moang*, PON *moahng head*, PTK *maŋo *forehead, top of head*. In contrast, none of the other CM forms cited have glosses that refer to the head: AA *ma-mano-na stomach, belly*, *mano-ara*, *mano-asa*, *mano-mano breathe*, *mano-na breast, chest; breath, respiration*; ARS *ma-maŋo-na pit of chest, stomach, where breath heaves*; KWO *maŋo breathe, pause for breath*, *fa'a-maŋo wait, take a break*, *maŋo-na breath, life*; LAU *maŋo breathe*, *maŋo-la to rest, have a spell*, *ata-maŋo to sigh, a sigh*; SAA *ma-maŋo breath*; (metaphorical) *heart*, (ma)maŋo *breathe*; ULW *maŋo-maŋo breathe*. While it is true, as Blust states (1984:111), that KIR mangongo has been reported with the meaning

passage between nose and throat; the breath, or rather its odour from the nose (Bingham 1908), neither this specific form nor any other form with a gloss pertaining to 'breathing' is attested elsewhere in Micronesia. Thus, the proposed comparison is quite tenuous.

Additionally, there is the possibility that Lakalai *mago sea*; *nasal mucus*; *brain* may reflect an earlier *maŋo which had a gloss which could have been independently limited in MC languages to the head and in CM languages to breathing (Chowning, personal communication). Chowning (1973) has observed other cases where Lakalai *g* reflects earlier *ŋ. Also, as Blust observes, Mono-Alu *mao-aha breathe* may reflect an earlier *maŋo.

(4) PMMC *lama lake, lagoon

Blust claims that this reconstruction represents a shared shift in the meaning of POC *laman *sea beyond the reef, deep blue sea*, and that the shift is related to the development of *mansawa discussed above (1984:115-117). His supporting data are AA *ramo lake*, ARS *rama water between reef and shore; strait between mainland and small island; long deep channel in the open sea*, ramarama *deep water beyond the edge of the reef*, LAU *lama pool at low tide in the reef*, SAA *lama lake*, KIR *nama lagoon, lake*, and MOK *lam lagoon*. To Blust's MC data may be added MRS *lam* and CRL *lama-*, both with the meaning *bay*.

My comments on this comparison are limited to noting that the ARS forms and those for MRS and CRL appear to continue the POC reference to *deep water* (albeit water which is relatively close to shore), and to observing, with Blust, that reflexes of POC *namo *sea within the reef, lagoon, deep pools near the shore* are well attested in both CM and MC languages: ARS *namo landlocked shallow lagoon near the shore*, LAU *namo the lagoon inside a reef, near the reef (the deep) pools towards the shore*, PMC *namwo *lagoon*, with the following reflexes: KIR *namo harbour*, MRS *naŋ secondary lagoon*, KSR *nwem deep area in lagoon*, PON *nahmw lagoon, deep place inside barrier reef*, PTK *namwo *lagoon* (with all reflexes showing that meaning). Thus, it would appear probable that *lama had not in fact replaced *namo in the meaning *lagoon* in the ancestors of either the MC or CM languages. It is likely, however, that both groups of languages have undergone a limiting of the meaning of *lama to something like *deep water close to the shore*.

(5) PMMC *mami fresh water

Except for the Ponapeic languages, which reflect a possibly innovative *pili *fresh water*, all MC languages reflect POC *(n)danum in this meaning. Since the only MC form cited by Blust for his reconstruction is KIR *mam fresh water* (cf. KIR *ran water, fresh water, liquid, juice, sap, milk* < POC *(n)danum), and he also cites only two CM forms (ARS *mami fresh, good drink, water; mixed with fresh water, as the sea near a river mouth*; KWO *maa-mami-la brackish water*), it would seem more likely that the KIR and ARS forms reflect somewhat parallel but independent innovations.

(6) PMMC *(m)puRo(m)puRo foam, bubbles

KIR *puropuro ebullition, bubbling up, frothing* does not, as Blust states, reflect this putative reconstruction, but rather PEO *(m)puso *foam, bubbles* (which is also reflected by PON *pwudopwud foam, scum*). Of the other MC forms cited by Blust, it is very probable that all but one reflect his own Proto-Malayo-Polynesian reconstruction *bureq, as Blust recognises. The remaining MC form, PON *pwolol to bubble, bubble*, is inconsistent with either earlier *R or *d, and thus more probably indicates a loan from the regular KSR reflex *pulohl blister*,

bubble (of hot water). While the provenance of the two CM forms cited by Blust is not clear (KWO bulobulo *eddy in stream, whirlpool*, LAU fulofulo *eddy in the sea when a turtle dives or a ship sinks*), it is very probable that they are unrelated to any of the MC forms.

(7) PMMC *pwela (*taro*) *swamp*

The MC forms cited by Blust all reflect a PMC *pwelu *dirt, soil, taro swamp*, which appears to have cognates in the Banks Islands in the type ^kpwolo *dirty* cited by Tryon (1976), and perhaps, although with a different initial consonant grade, in Fijian duka-veluvelu *very dirty*. The only CM form cited by Blust, ARS bwera *swamp*, has a final vowel which cannot be reconciled with the MC forms.

(8) PMMC *ŋ(i,u)so *squid, cuttlefish*

Blust proposes this reconstruction as reflecting a phonological innovation by which the initial alveolar nasal of POC *nu(n)s(i,o) became a velar nasal in LAU and in some MC languages. At best, it is a highly problematic comparison, as AA, ARS, KWO, and SAA all reflect only the alveolar nasal, and LAU has the expected form nuto *squid* in addition to the innovative ŋuto, in the same meaning. Within MC, MOK and PON nuhd, MRS nōt also point to a PMC *nu(d,z)o *squid*, with an alveolar nasal. In fact, only the Trukic languages of Micronesia consistently reflect a velar nasal in this form: PTK *ŋudi or *ŋido > TRK niiti-, MRT ngúút, PUL ngiito-, STW ngúút, CRL ngúti-, WOL ngiit0, ULI ngidi-, PUA ngiit0. Since the Trukic languages were probably the last group to break off in the dispersal of the Proto-Micronesian community (Jackson 1983), there would appear to be few grounds for concluding that the velar nasals in LAU and PTK resulted from common inheritance of a velar nasal in an ancestral language. Blust himself regards the comparison as a "difficult" one (1984:121), and given the subgrouping relationships now held to be true of Micronesia, it is almost certain that the LAU and TK forms reflect independent developments, perhaps in both cases as a result of contamination from languages reflecting PPN *ŋuu *squid, cuttlefish*.

I have nothing to add at this time to Blust's other proposed comparisons. However, it is clear that Blust himself considers some of the comparisons which we have been discussing to be the strongest evidence for the proposed MMC grouping.³² He writes (1984:132), "The evidence for the MMC hypothesis is limited, but includes several almost certain innovations in basic vocabulary (*nose, sea, breathe*) and at least one apparent semantic drag chain [that involving putative PMMC *mansawa and *lama]." As we have seen, the innovative status of these forms is not as certain as Blust believes; as a result, the search for conclusive evidence of the external relationships of the Micronesian languages must continue.

3.3 The proposed linkage between the Admiralty Islands and Micronesia

Smythe (1970) proposes a possible connection between languages of the Admiralty Islands and those of Micronesia on the basis of 146 examples of lexical similarities between various Admiralty languages and Trukese and of some grammatical similarities between the two. Because the putative grammatical similarities are not described in a fashion which permits systematic comparison, we are left to examine the proposed lexical comparisons. Those, in turn, have been strongly criticised by Blust (1984) on the grounds that only five of the proposed comparisons "exemplify recurrent phonological correspondences and do not reflect

an established comparison " (1984:128). Blust goes on to demonstrate that even those five comparisons should be discarded on various grounds,³³ and concludes that "Smythe's argument for an Admiralty-Nuclear Micronesian subgroup rests on virtually no solid evidence of exclusively shared lexical innovations" (1984:129).

Before rejecting Smythe's proposal out of hand, however, it may be of value to note a few of Smythe's other proposed comparisons, for, although each is problematic in its own way, they may together indicate the need for further research into the matter. The comparisons in question are as follows:

(1) Gele' ba:bi *beach*, Seimat helpi *sand*, "others" pi:a *sand*, TRK pi *sand* (PMC *pi(k,x)a *sand, sandbank, beach*). Blust (personal communication) has also brought Wuvulu pie *sand, earth* to my attention, but suggests that it and the other Admiralty forms reflect POC *(m)pia *earth, ground, land*. However, the fact that only the Admiralty forms share the meaning *sand* with MC languages, and the additional fact that Wuvulu, at least, apparently loses earlier *k and *ŋk (> PMC *k and *x, respectively) may suggest a connection between the Admiralties and Micronesia.

(2) Gele' ne-m̄lulu, Bipi pulu, Loniu ñapulu-, TRK punuwa- *spouse* (PTK *pulua *spouse, to marry*). Blust (personal communication) states that this form is of "rather limited distribution in the Admiralties" and that it does not reflect the final vowel of the Trukic reconstruction. The comparison may be worth noting, nonetheless.

(3) Gele' n̄amat, Lou ramat, TRK aramas *person* (PMC *aramata). Ross (1977) reconstructs Proto-Siau *Ramat(a), Proto-Kairiru *Ramat *person*, which may be cognate with the MC and Admiralty forms. However, Siau, Kairiru *R is reported by Ross as normally reflecting POC *R, which is apparently lost regularly in the Admiralties (Pawley 1977). It becomes possible, thus, that either the MC and Admiralties languages instead reflect earlier *(n)d or that both groups reflect a merger of earlier *R and *(n)d.³⁴ In either case, there may be evidence here of a connection between the two groups.

(4) Seimat i:a, TRK iya (PMC *ia(a)) *where?* Blust (personal communication) has also directed my attention to Wuvulu ia, in the same meaning. As observed earlier, this etymon also seems to be attested in the southern New Hebrides and in Nauruan. I have not found other examples, however, so its rare occurrence in the Admiralties and Micronesia may provide some support for linking those areas.

Additionally, Blust (personal communication) has reconstructed Proto-Eastern Admiralties *kandahV *cloud*, which is almost certainly cognate with PMC *kaca(w)u *cloud, raincloud* but to my knowledge is not attested elsewhere. Blust states (personal communication) that PEAdm *-h- points to an earlier *p, but that is also compatible with the PMC reconstruction.³⁵

There is very little data published on the languages of the Admiralty Islands, and although Robert Blust has kindly shared with me some of his 200-word lists on a few Admiralty languages, I have no information on the grammars of those languages and almost none on less basic vocabulary. Under these circumstances, the rather weak comparisons just described may be grounds for investigating further the possibility of a connection between the Admiralties and Micronesia. The two groups are in reasonably close geographical proximity, and there also appears to be some non-linguistic evidence of a connection between them (Chowning 1977:102n).³⁶

4. CONCLUSION

Although some steps have been taken in this paper toward the formal establishment of a Greater Micronesian subgroup of Oceanic, to consist of Nauruan and the Micronesian languages, in one sense little else has been accomplished in the determination of the original source of the Micronesian languages. We have seen that Pawley's carefully researched criteria for including a language in the putative Eastern Oceanic group have failed to provide secure evidence for such a determination, as has Blust's detailed proposal for linking Micronesian languages with those of Cristobal and Malaita. On the other hand, Smythe's proposal, while justly criticised by Blust, may offer a few tantalising suggestions of a linkage of some sort between the Admiralty Islands and Micronesia.

Part of the problem may stem from the fact, now more widely recognised, that Micronesian languages are lexically quite conservative for Oceanic languages, despite their somewhat innovative phonologies. Robert Blust (personal communication) has recently recognised the following MC retention percentages for his Malayo-Polynesian 200-word list: KSR 28.6, MRS 29.9, KIR 32.0, PON 30.2, TRK 38.3, and PUA 37.9, and my own computations using the same list average 2-3 percentage points higher than Blust's. These high retention percentages are also implied by the shared cognate percentages of five Micronesian languages and 31 other Oceanic languages which I have computed on a modified 100-word list which was designed to highlight some typical Micronesian lexical distinctions, such as separate lexical entries for *arm* and *hand*, and also for *sea water*, *fresh water*, and *ocean*. These percentages, which are shown in Table 6, show that MC languages have their highest number of shared cognates with such geographically dispersed languages as Kapingamarangi, Fijian, Nggela, Longgu, Faghani, and Wuvulu, and their second highest percentages with Nukuoro, Rotuman, Raga, and Lakalai.

It must be emphasised, lest too much be made of these cognate percentages, that the great majority of shared cognates indicated in Table 6 are reflexes of clear POC reconstructions, and most, if not all, of the others reflect probable POC etyma. Many of these latter would appear to point to POC doublets with the same (or very similar) meanings. For example, PMC and several languages of the New Hebrides and Admiralties reflect a type *qalo *sun* instead of the more widely reflected *qanso. It would also seem possible on the basis of data already discussed to reconstruct for the meaning *where?* both POC *p(i,e)a and POC *pei, and quite possibly a POC *ia. Similarly, although the supporting data are more limited, it would appear necessary at present to reconstruct both POC *ŋaqija and POC *ñaqija for the meaning *when?*, and both POC *(n)sai and *ta(i) for the meaning *who?*

In fact, of course, it is extremely unlikely that these and other well known apparent doublets had exactly the same meaning in Proto-Oceanic. It is much more likely that, if both forms were in fact present in the proto-language, there was a semantic distinction between them which has since become obscured by varying processes of historical change. To gain a better understanding of the grammar and lexicon of Proto-Oceanic, one of our tasks will be to disambiguate the meanings of such forms. That task has been begun in much of the recent work of Robert Blust and in Andrew Pawley's analysis of POC terms for *people* (Pawley 1979b), but there is still a very great deal of work which needs to be done in this area.

And this is only one of the areas relevant to the reconstruction of POC in which work is needed. The truth, I believe, is that we still know far too little about the lexicon, grammar, or even phonological developments of POC to identify the shared innovations needed for us to propose higher-order subgrouping relationships

Table 6: Cognate percentages of Micronesian and other Oceanic languages

	KSR	KIR	MRS	PON	WOL
1. Rotuman	27	25	28	27	26
2. Fijian	26	26	28	26	37
3. Polynesian languages					
Tongan	24	23	27	24	28
Kapingamarangi	30	26	32	28	33
Nukuoro	25	24	27	28	29
Rennellese	24	22	24	25	26
Ontong Javanese	28	21	26	24	28
Futunan	21	21	23	23	25
4. NH languages					
Mota	22	24	25	22	30
Mosina	20	21	19	18	24
Raga	23	25	26	28	35
Nokuku	18	18	21	18	25
Nambel	19	20	21	20	24
Uripiv	22	19	25	22	28
Sesake	22	28	19	19	26
Nguna	23	27	20	20	27
5. South-east Solomons languages					
Bugotu (Dhadaje)	21	22	25	24	24
Nggela	24	25	31	29	33
Longgu	26	29	29	28	29
Lau	22	26	27	26	25
Kwaio	25	25	23	23	24
'Are'are	20	23	23	25	26
Sa'a	18	23	22	23	24
Arosi	23	27	25	26	25
Faghani	25	29	27	28	29
6. Admiralty Islands languages					
Wuvulu	29	27	31	31	36
Loniu	21	21	23	21	25
Lou	23	24	24	22	22
7. Other Oceanic languages					
Roviana	15	17	17	16	16
Motu	24	23	21	24	26
Lakalai	28	23	25	26	31

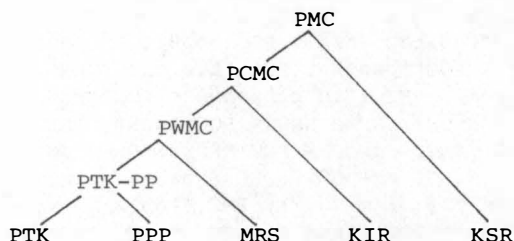
that will withstand intensive examination. In order for us to securely identify such relationships, we need first to clearly define more lower-order groups like those of Polynesia, Central Pacific, and (I believe) Micronesia, and to take major steps toward the reconstruction of the proto-languages of those subgroups. We need a fuller understanding of phonological developments within Oceanic, including, for example, the extent to which phonological changes affecting *R and other POC segments may have been diffusing through the lexicons of Oceanic languages. We need further work toward the disambiguation of apparent POC doubts like those described above. We also, I submit, need further formal attempts at identifying higher-order subgrouping relationships like those of Pawley and Blust which have been discussed in this paper, for even if such attempts, when put to the test, fail to provide the conclusive evidence which is needed, the testing of them teaches us not only about the languages involved, but also about the nature of Proto-Oceanic.

NOTES

- I am indebted to Robert Blust, Ann Chowning, and Paul Geraghty for their comments and suggestions on an earlier draft of this paper. I retain, however, full responsibility for any errors of fact or interpretation which remain.
- Abbreviations used in Table 1 are as follows:

CRL	Saipan Carolinian	PPP	Proto-Ponapeic
KIR	Kiribati (Gilbertese)	PTK	Proto-Trukic
KSR	Kosraean (Kusaiean)	PTK-PP	Proto-Trukic-Ponapeic
MAP	Old Mapian	PUA	Pulo Anna
MOK	Mokilese	PUL	Puluwatese
MRS	Marshallese	PWMC	Proto-Western Micronesian
MRT	Mortlockese	SNS	Sonsorolese
PCMC	Proto-Central Micronesian	STW	Satawalese
PMC	Proto-Micronesian	TRK	Lagoon Trukese
POC	Proto-Oceanic	ULI	Ulithian
PON	Ponapean	WOL	Woleaian

The internal subgrouping hypothesis for Micronesian that is assumed in this paper is as shown below (see Jackson 1983 for supporting evidence):



Much of the Micronesian data on which this analysis is based is in Bender et al (in preparation). Other data have been taken from the sources provided in the References, except for those for CRL, MRT, and STW, which are from my own fieldnotes.

- The standard orthography of the source for each language is used in this table. Footnotes provide the phonetic values of symbols which may be ambiguous.
- Labiovelar *pw is reconstructed before a non-round vowel for the following Micronesian etyma: PMC *pwaŋa *hole, cave*; PMC *pwelu *dirt, soil; dirty*; PMC *pwake(a) *hawksbill turtle*; PMC *pwara *pubic area*; PMC *pwa 'future aspect'; PMC *pwauzu *nose*; PMC *pwa *say, tell*; PWMC *(r,c)epwa *beard*; PMC *pwaki *lift, carry, take*; PMC *pwexa *twins*; PMC *pwau *pound food*; PMC *pwau *fishing pole*; PCMC *pwaro *bent, curved*; PMC *pwapwu *shark species*; PWMC *pwalu *to cover (s.t.)*; PCMC *pwaipwai *foolish, stupid*; PCMC *pwa *because*; PWMC *pwadu *scar*; PMC *pwace *coral lime*; PMC *pweci *hot*; PMC *lipwa *pit, hole*; PWMC *pwalu *taro patch, dirt*; PTK *pwariku *kind of dance*; PTK *pwaro *box, crate, container*; PTK *ka-pwata *shout, call*; PTK *pwadai *fat, obese, ripe*; PTK *dapwa *follow*; PTK *kupwa *foot and leg, footprint*.
- Labiovelar *mw is reconstructed before a non-round vowel for the following Micronesian etyma: PMC *mwaau *good, seemly, fitting*; PMC *umwa *hermit crab*; PCMC *mweŋau *eat; food*; PMC *mwata *earthworm*; PMC *mware *lei, garland*; PWMC

*mwamwane *able, capable, good*; PMC *mwaane *man, male*; PCMC *mwalu *joint (as elbow)*; PMC *mwakumwaku *arrowroot, starch*; PMC *mwakuzu *move, shake*; PWMC *mwak(i,a) *needlefish*; PMC *mwacako *steal, act greedily*; PMC *(u,i)mwa *house, home*; PMC *camwa *forehead, brow; housegable*; PMC *mwane *cross-sibling*; PTK *mwan̄ea *cross-sibling*; PTK-PP *mwakarikari *the Pleiades*; PTK *amwa *struts connecting outrigger float to boom*; PTK-PP *mwetu *grant, allow; adopt child*; PTK *cimwa *head, bundle*; PTK *mwano *secret, hidden, shaded, lost*; PTK *mwaku *Pisonia grandis*; PTK *mwacani *want, desire, agree to*; PTK *ka-mwacu *hold, grasp*; PNTK *mwiici *meeting, to meet*; PNTK *mweli *main sheet for canoe boom*; PNTK *mweika *chili pepper*; PNTK *kamwee *large tridachma*; PNTK *ka-imwainwa *shelter on lee platform of sailing canoe*.

6. POC *R is reflected as PMC *r in the following etyma: PMC *mera *red, stained* < POC *meRa; PMC *afara *shoulder* < POC *(qa)paRa; PMC *daku-laara *swordfish, marlin* < POC *laya(R) *sail*; PMC *maraara *light in weight* < POC *maRaqa; PMC *luru *shade, shady* < POC *ma-luR; PMC *-ŋaratu 'counting classifier for units of thousands' < POC *Ratu(s) *hundred*; PMC *para(ta) *tradewind, windstorm* < POC *(m)paRa(ta) *north-west monsoon*; PMC *ropa *smashed, broken* < POC *Ropa; PMC *takuru *back* < POC *takuRu; PMC *taŋiri *tuna species* < POC *taŋiRi; PMC *wakara *root* < POC *waka(Ri); PMC *pwoRo *wring, squeeze* < POC *(m)poRo. In contrast, POC *R is lost in the following PMC forms: PMC *pau *arm, wing* < POC *(m)paRu (Blust, personal communication); PMC *(ca)caa *blood* < POC *(n)daRa; PMC *cui *bone* < POC *nduRi (Blust 1978); PMC *dakau *reef, reef island, atoll* < POC *nsa(ŋ)kaRu *reef, coral*; PMC *dou *spear, stab, inject* < POC *(n)saRu; PMC *fai *ray fish* < POC *paRi; PMC *faka-afi *evening* < POC *Rapi; PMC *fau *new* < POC *paqoRu; PMC *kia *mat* < PEO *kiRe (Geraghty, personal communication); PMC *fau *hibiscus* < POC *paRu; PMC *kuita *octopus* < POC *kuRita; PMC *ñañewa *yesterday* < POC *ñoRa; PMC *(u,i)mwa *house* < POC *Ruŋma; PMC *uda *load, cargo* < POC *RuJa; PMC *ua *neck* < POC *Ruqa; PMC *taui *conch, trumpet* < POC *tapuRi; PMC *ŋaŋaa *duck, booby* < POC *ŋaRa (cf. Roviana ngara, Fijian gaa *wild duck*); PMC *tapia *bowl, dish* < POC *tampiRa (cf. Manam tabira *dish*); PMC *auda *current* < POC *qaRusa; PMC *pwauzu *nose* < earlier *Ru(n)su (see Section 3.2 for discussion of the last form).
7. Paul Geraghty's suggested PEO *z is used as the base of comparison here because of the fact that it corresponds in a very regular fashion with PMC *z (Geraghty 1983:130-148; also see Jackson 1983:343-350 for detailed discussion of the MC reflexes). It should also be noted that Geraghty (1983: 90-91 and personal communication) believes that his PEO *z, which corresponds to reconstructed POC *ns, in fact reflects an oral grade palatal, while his PEO *s reflects the nasal grade. If Geraghty is correct, then the PMC merger of *z and *j in fact involves the merger of two oral grade segments, while the PMC merger of PEO *s and *nj (7) involves two nasal grade segments.
8. Geraghty (1983) writes this reconstructed segment as *j, but he has since agreed to reserve that symbol for the POC reconstruction made by Blust (1978) and to use *nj for his own PEO reconstruction (Geraghty, personal communication).
9. In addition, it appears very likely that the following processes affecting vowels can be reconstructed for PMC: (1) shortening of all word-final vowels, and loss of word-final *i after sonorant consonants; (2) regressive vowel assimilation affecting adjacent vowels and also those separated by a single consonant; (3) lengthening of the first syllable in a bimoric phrase so that each phrase has at least three mora (see Rehg 1984 for formalisation

of this process and discussion). It is also possible, as Marck (1977) has suggested, that phonetic qualities of the consonants may have affected neighbouring vowels.

10. A cursory scanning of the subject, object, and focus pronoun sets for other Oceanic languages indicates that the problems discussed for MC are not limited to that group of languages. For example, of the 30 languages treated in Pawley 1972, 21 have identical focus and object pronouns in the meaning '1pl.exc.', and 11 have identical focus and subject pronouns in the same meaning. In addition, Pawley (1972:64) reports six instances of alternate forms among the subject pronouns. For the '2pl.' forms, Pawley (1972:66) reports 24 instances of identical forms in the focus and object pronouns and nine instances of identical focus and subject pronouns, with seven cases of alternate subject pronouns. It appears clear that considerable work still needs to be done in the reconstruction of POC personal pronouns.
11. Other such languages include Ulawa and Marau Sound.
12. The Trukic, Ponapeic, and Marshallese forms for the meaning *hither, towards speaker* all reflect PEO *(n)soko *arrive*. That reconstruction has cognates throughout Micronesia in the meaning *come, arrive* (PMC *doko), but only in the Western Micronesian languages is its reflex also used as a directional enclitic. The fact that a reflex of PMC *doko has replaced earlier *mai in this function in those languages is powerful evidence for subgrouping them together.
13. It may be noted at this time that the Rotuman directional morphemes correspond almost exactly with those reconstructed for PMC (Churchward 1940:41):

	<i>thither</i>	<i>outwards</i>	<i>down</i>	<i>up</i>	<i>hither</i>	<i>inwards</i>	<i>away</i>
PMC	*watu	*otu	*zio	*zake	*mai	*loŋo	*lako
ROT	afu	hofu	sio	seʔe	me	loga	(age)

Only the items which are glossed *away* are not cognate.
14. KSR has a form la- in this meaning. Its source is not known.
15. The Trukic and Ponapeic languages have reflexes of the type *kana- as classifiers for edible objects. This type is cognate with the Kiribati transitive verb kana *to eat*, and reflects POC *kani *to eat*. However, the verb for *to eat* among Western Micronesian languages is *kaŋi, reflecting an innovation which also occurs in KSR and in some of the forms of the paradigm in KIR.
16. MRS has kije- in this meaning, apparently reflecting an earlier *kita.
17. Loss of the final syllable of POC *pati *four* has also occurred in Fijian and Polynesian (Pawley 1972) and in Lakalai (Chowning 1973).
18. Although PMC *-cau 'counting classifier for thin flat objects' is only reflected in Trukic and Ponapean within Micronesia, the fact that cognates occur in the same function and meaning in Nauruan (Nathan n.d.) and Lau Fijian (Geraghty, personal communication) requires its reconstruction for Proto-Micronesian.
19. Lynch and Tryon (1985) state that such forms "show irregular loss of the initial *p [of *pia *where?*]", but it would seem far more probable that they are regular cognates of an earlier *ia which is also reflected in Micronesia.
20. Nathan's work on Nauruan consists of a paper presented at the First Conference on Austronesian Linguistics (Nathan 1973), a very brief unpublished

description of some aspects of Nauruan grammar (Nathan n.d.a), and an unpublished short word list keyed to Grace's Proto-Oceanic finderlist (Grace 1969). All of my data on Nauruan are taken from these sources.

21. It appears certain that NAU has borrowed heavily, perhaps from several languages. A clear source for some of the borrowings is Kiribati, but other possible sources have not been identified. Also not yet determined are the periods when borrowing may have occurred.
22. Blust (1984:109) cites NAU *bodi-n nose* as a possible cognate of his proposed PMMC *pwaRusu *nose*, but states that the NAU form "fails to agree with the sound changes suggested by Nathan (1973)". Nathan (1973:492) had tentatively concluded that NAU reflects *R as either r or ʀ, but, as we have seen, the most frequent NAU reflex of POC *R is Ø. In this as in all other respects, NAU *bodi-n* (more correctly *bwoodi-n*) is an entirely regular reflex of Blust's proposed reconstruction.
23. Although it is quite clear now that Yapese is not a "Micronesian language" as that term is defined in section 1 of the present paper, it is not at all certain what group Yapese does belong to. It is not even conclusive that the language is Oceanic (see Bradshaw 1975).
24. Pawley (1979a) shows that Rotuman should also probably be included within this group and that Rotuman, Fijian, and Polynesian form a Central Pacific grouping.
25. Pawley suggested in 1972 that the fact that Kiribati loses *R might be grounds for its inclusion in the proposed HC. Since he considered Kiribati as representative of the Micronesian languages and was also aware of the retention of some instances of *R in at least the Trukic languages, however, it is not clear why he emphasised the Kiribati development.
26. The fact that both MC and NH languages reflect initial syllable reduplication may be noteworthy, although the same pattern of reduplication is attested in, e.g., Sa'a *nonola* and Fijian *nanoa yesterday*.
27. According to Tryon (1976), Mota *n* is a regular reflex of POC *nd.
28. Geraghty (1978) suggests on the basis of data from the Shepherd Islands that the form *(n)daqe *blood* be reconstructed in competition with *(n)daRa. If Geraghty is right, some of these forms may reflect that reconstruction.
29. Pawley (1979:43n) states that the fact that many languages in the New Hebrides show variation between loss of *R and its merger with *d suggests "borrowing". Since most New Hebridean languages now appear to exhibit such variation, it is difficult to imagine the scale of such borrowing, or how, if Pawley's suggestion is correct, it would ever be possible to make firm subgrouping decisions from New Hebridean data.
30. The fact that South-east Solomonian languages and apparently some New Guinea languages attest the merger of *R and *l may suggest a third competing change. If so, that change might have been the cause of the problematic NH reflexes of POC *Ruqa *neck* discussed above.
31. The fact that Bugotu *maha* does not appear to have a long final vowel would argue against this possibility. There is some evidence that Bugotu shortens some long vowels, however: cf. Bugotu *mane man, male person* < POC *ŋmaqane.
32. Blust (personal communication) has recently stated that he considers the proposed reconstruction of PMMC *tina *mate orphan* to represent a significant

morphological innovation of the proposed group. The form is attested in AA, ARS, LAU and SAA within CM, but only in MRS and Woleaian within MC. Its significance appears somewhat questionable, however, on two grounds: (1) the transparency of the form's meaning makes independent development likely, and (2) forms with the meaning *orphan* are not regularly provided in published sources on Oceanic languages. For example, Tryon 1976, which Blust used as his source for many of the New Hebridean languages which he checked, does not include forms for *orphan*.

33. The five comparisons in question are as follows:

- (1) Gele' and "others" mwan *bad*, TRK mwän *error*;
- (2) Gele' and "others" wasiw *cane*, TRK was *stem (grass)*;
- (3) Gele' os, TRK ot *taro*;
- (4) Gele', TRK ηat *hole*; and
- (5) Gele' ηaw *cat*, TRK ηäo *mew*.

Of the five cited TRK forms, only (3) woot *taro* reflects a secure PMC reconstruction: PMC *odo *taro* < POC *oso *journey provisions*. Blust, however, states that his own fieldnotes fail to show any form for *taro* similar to os in Gele' or any other Admiralty language and discards the comparison on that basis. TRK mwmwáán *to err, make a mistake, be wrong, mistaken, incorrect, to be done improperly or without permission* might correspond with Gele' mwan *bad*, but does not appear to be attested elsewhere in MC. TRK ngaat *hole* has cognates throughout Trukic, but not in other MC languages. Moreover, the reconstruction of the medial consonant is somewhat problematic for PTK: *ηa(t,d)a *hole*. As Blust suggests, the comparison in (5) can probably be rejected as parallel onomatopoeic developments. In any case, cats were only very recently introduced to Micronesia. Blust is also correct in noting that the final vowels of Gele' wasiw *rattan* and TRK wasa- *handle (as of axe or basket), shaft (of spear), pole (for flag)* are irreconcilable.

34. The MC languages could only reflect earlier *d or *R.

35. Paul Geraghty (personal communication) also points out that the Fijian island name Kadavu corresponds formally in perfect fashion with the MC and Admiralties forms. It is not at all clear whether there is an etymological connection.

36. Chowning (1977:102n) writes, "The most westerly islands of the Bismarck archipelago, especially Aua and Wuvulu, contain people whose physical appearance and technology indicate, at the very least, strong Micronesian influence" Technological similarities include those of gardening techniques and the appearance of the canoes (Chowning, personal communication).

REFERENCES

- ABO, Takaji, Byron W. BENDER, Alfred CAPELLE, and Tony DeBRUM
 1976 *Marshallese-English dictionary*. Honolulu: University Press of Hawaii.
- BENDER, Byron W.
 1971 *Micronesian languages*. In T.A. Sebeok, ed. *Current trends in linguistics*, vol.8, 426-465. The Hague: Mouton.

BENDER, Byron W., ed.

1984 *Studies in Micronesian linguistics*. PL, C-80.

BENDER, Byron W, Robert W. HSU, Frederick H. JACKSON, Kenneth L. REHG, Stephen TRUSSEL and Judith W. WANG

in Micronesian cognate sets. Computer printout. Department of prep. Linguistics, University of Hawaii.

BIGGS, Bruce

1979 Proto-Polynesian word-list II. MS.

BINGHAM, Hiram

1908 *Gilbertese-English dictionary*. Cambridge, Mass.: The University Press.

BLUST, Robert A.

1978 *The Proto-Oceanic palatals*. The Polynesian Society, Memoir No.43.

1984 Malaita-Micronesian: an Eastern Oceanic subgroup? *Journal of the Polynesian Society* 93/2:99-140.

BRADSHAW, Joel

1975 Problems in determining the relationship of Yapese to Proto-Oceanic. MS.

1978 Notes on subgrouping in the Huon Gulf area. *Working Papers in Linguistics, University of Hawaii* 10/1:49-82.

CAPELL, Arthur

1969 *Grammar and vocabulary of the language of Sonsorol-Tobi*. *Oceania Linguistic Monographs* 12.

CHCWINING, Ann

1973 Milke's "New Guinea cluster": the evidence from north-west New Britain. *Oceanic Linguistics* 12/1-2:189-244.

1977 *An introduction to the peoples and cultures of Melanesia*. 2nd edn. New York: Cummings.

CHURCHWARD, C. Maxwell

1940 *Rotuman grammar and dictionary*. Sydney: Australasian Medical Publishing Co.

CODRINGTON, Robert H.

1885 *The Melanesian languages*. Oxford: Clarendon Press.

DEMPWOLFF, Otto

n.d. Grammar of the Graged language. Lutheran Mission, Narer, Karkar Island. Mimeo.

DYEN, Isidore

1949 On the history of the Trukese vowels. *Language* 25:420-436.

ELBERT, Samuel H.

1947 Ulithi-English and English-Ulithi word list. United States Naval Military Government. Mimeo.

- 1972 *Puluwat dictionary*. PL, C-24.
- 1974 *Puluwat grammar*. PL, B-29.
- FRITZ, Georg
- 1911 *Die Zentralkarolinische Sprache*. Berlin: Georg Reimer.
- GERAGHTY, Paul
- 1978 Review of Tryon 1976. *Kivung* 11/1:75-86.
- 1983 *The history of the Fijian languages*. *Oceanic Linguistics* Special Publication No.19.
- GOODENOUGH, Ward H. and Hiroshi SUGITA
- 1980 *Trukese-English dictionary*. American Philosophical Society.
- GRACE, George W.
- 1955 Subgrouping of Malayo-Polynesian: a report of tentative findings. *American Anthropologist* 57/2:337-339.
- 1964 Movement of the Malayo-Polynesians: 1500 B.C. to A.D. 500: the linguistic evidence. *Current Anthropology* 5/5:361-368.
- 1969 A Proto-Oceanic finder-list. *Working Papers in Linguistics, University of Hawaii*, 1/2:39-84.
- 1971 Tentative Oceanic subgrouping. In Grace 1981.
- 1976 Review Article: Green and Kelly 1972. *Journal of the Polynesian Society* 85/1:103-112.
- 1981 Tentative Oceanic subgrouping. MS.
- GRACE, George W., James A. THARP, Peter C. LINCOLN and Janet BLACK
- 1984a Proto-Oceanic cognate sets. Computer printout. Department of Linguistics, University of Hawaii.
- 1984b Proto-Oceanic finder-list. Computer printout. Department of Linguistics, University of Hawaii.
- GREEN, R.C. and M. KELLY, eds
- 1972 *Studies in Oceanic culture history*, vol.3. Pacific Anthropological Records 13. Honolulu: Department of Anthropology, Bernice P. Bishop Museum.
- GROVES, Terab'ata, Gordon W. GROVES and Roderick A. JACOBS
- 1985 Kiribatese: an outline grammar and vocabulary. PL, D-64.
- HARRISON, Sheldon P.
- 1976 *Mokilese reference grammar*. Honolulu: University Press of Hawaii.
- 1978 Transitive marking in Micronesian languages. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, Fascicle 2: *Eastern Austronesian*, 1067-1127. PL, C-61.
- 1981 Where the Micronesian possessive classifiers might come from, and why Gilbertese doesn't have any. Paper presented at a meeting of the Austronesian Circle, University of Hawaii. September 3, 1981.

- n.d. An indirect argument for a Proto-Micronesian article. MS.
- HARRISON, Sheldon P. and Salich ALBERT
- 1977 *Mokilese-English dictionary*. Honolulu: University Press of Hawaii.
- HARRISON, Sheldon P. and Frederick H. JACKSON
- 1984 Higher numerals in several Micronesian languages. In Bender, ed. 1984:61-79.
- IVENS, W.G.
- 1933 A grammar of the language of Bugotu, Ysabel Island, Solomon Islands. *Bulletin of the School of Oriental and African Studies* 7:141-177.
- 1937 A grammar of the language of Florida, British Solomon Islands. *Bulletin of the School of Oriental and African Studies* 8:1075-1110.
- 1940 *A dictionary of the language of Bugotu*. London: The Royal Asiatic Society.
- JACKSON, Frederick H.
- 1983 The internal and external relationships of the Trukic languages of Micronesia. Ph.D. dissertation, University of Hawaii.
- 1984 On lexical diffusion as subgrouping evidence. Paper presented at the annual meeting of the Linguistic Society of America, Baltimore, Maryland. December 30, 1984.
- JACKSON, Frederick H. and Jeffrey C. MARCK, compilers
- forth- *Carolinian-English dictionary*. Honolulu: University Press of coming Hawaii.
- JENSEN, John Thayer
- 1977 *Yapese-English dictionary*. Honolulu: University Press of Hawaii.
- KĀ'EO, Karen
- n.d. Eastern Oceanic and the nuclear Micronesian languages. MS.
- LEE, Kee-dong
- 1975 *Kusaiean reference grammar*. Honolulu: University Press of Hawaii.
- 1976 *Kusaiean-English dictionary*. Honolulu: University Press of Hawaii.
- LEVY, Richard
- n.d. Languages of the Southeast Solomon Islands and the reconstruction of Proto Eastern Oceanic. MS.
- LICHTENBERK, Frantisek
- 1977 Third palatal reflex in Manam. Mimeo from class presentation.
- LYNCH, John and D.T. TRYON
- 1985 Central Eastern Oceanic: a subgrouping hypothesis. In Andrew Pawley and Lois Carrington, eds *Austronesian linguistics at the 15th Pacific Science Congress*, 31-52. PL, C-88.
- MARCK, Jeffrey C.
- 1977 A preliminary manual of Micronesian sound correspondences. Paper presented at the Symposium on Austronesian Linguistics at the

Summer Meeting of the Linguistics Society of America, Honolulu, Hawaii, August 18-20, 1977.

NATHAN, Geoffrey S.

- 1973 Nauruan in the Austronesian Language Family. *Oceanic Linguistics* 12/1-2:479-501.
- n.d.a A grammatical sketch of Nauruan. MS.
- n.d.b [Nauruan word-list]. MS.

ODA, Sachiko

- 1977 The syntax of Pulo Annian: a nuclear Micronesian language. Ph.D. dissertation, University of Hawaii.

PAWLEY, Andrew K.

- 1972 On the internal relationships of Eastern Oceanic languages. In Green and Kelly, eds 1972:1-142.
- 1977 On redefining "Eastern Oceanic". MS.
- 1979a New evidence on the position of Rotuman. MS.
- 1979b Proto-Oceanic terms for people: a problem in semantic reconstruction. MS.

QUACKENBUSH, Edward Miller

- 1968 From Sonsorol to Truk: a dialect chain. Ph.D. dissertation, University of Michigan.

REHG, Kenneth L.

- 1981 *Ponapean reference grammar*. Honolulu: University Press of Hawaii.
- 1984 The origins of "compensatory lengthening" rules in Micronesian languages. In Bender, ed. 1984:53-59.

REHG, Kenneth L. and Damian G. SOHL

- 1979 *Ponapean-English dictionary*. Honolulu: University Press of Hawaii.

ROSS, Malcolm

- 1977 Relationships of the Austronesian languages of the Sepik and western Madang coast of New Guinea. Mimeo.
- 1979 The Austronesian languages of Papua: towards a family tree. MS.

SABATIER, Fr. E.

- 1971 *Gilbertese-English dictionary*. (Translated by Sr Oliva). Sydney: South Pacific Commission Publications Bureau.

SMYTHE, W.E.

- 1970 Melanesian, Micronesian, and Indonesian features in languages of the Admiralty Islands. (Edited by Alan Healey). In S.A. Wurm and D.C. Laycock, eds *Pacific linguistic studies in honour of Arthur Capell*, 1209-1234. PL, C-13.

SOHN, Ho-Min

- 1975 *Woleaian reference grammar*. Honolulu: University Press of Hawaii.

SOHN, Ho-Min and Byron W. BENDER

1973 *A Ulithian grammar.* PL, C-27.

SOHN, Ho-Min and Anthony TAWERILMANG

1976 *Woleaian-English dictionary.* Honolulu: University Press of Hawaii.

SUGITA, Hiroshi

n.d. *Trukese reference grammar.* MS.

TRYON, D.T.

1976 *New Hebrides languages: an internal classification.* PL, C-50.

TRYON, D.T. and B.D. HACKMAN

1983 *Solomon Islands languages: an internal classification.* PL, C-72.

WANG, William S-Y.

1969 *Competing changes as a cause of residue.* *Language* 45/1:9-25.

1979 *Language change - a lexical perspective.* *Annual Review of Anthropology* 8:353-371.

STEM-INITIAL CONSONANT ALTERNATION IN THE LANGUAGES OF EPI, VANUATU: A CASE OF ASSIMILATION?

D.T. Tryon

1. INTRODUCTION

Early scholars of the Island Melanesian languages such as Codrington (1885) and Ray (1926) observed alternations between certain consonants in certain Vanuatu languages, but without assigning any function to them. Ivens too noted such consonantal alternations in later studies (Ivens 1937-39, 1939-42). Here again the basis of the alternations was not really understood.

More recently a number of scholars have turned their attention to the phenomenon of consonant alternation within Oceania as a whole, but particularly in the languages of Vanuatu, for example Tryon (1973), Lynch (1975), Walsh (1982), Crowley (1982) and Clark (1985).

Lynch, among these writers, was the only one to propose a detailed explanation of consonantal alternation within Oceanic languages as a whole, in a hypothesis which claimed that the alternations observed were the result of the assimilation of the 'realis' prefix *ma to the initial consonant of certain verb stems, this prefix being found almost throughout central and northern Vanuatu. Beyond Vanuatu he found that alternations also resulted from a fusion of the 'irrealis' marker *na with the initial consonant of verbs. This latter fusion was found by Lynch to be confined to the Papua New Guinea area, in the languages of Yabem and Mapos Buang in the Morobe District.

Until the present, however, there has been no comparative study of stem-initial consonant alternation across languages within a single subgroup, largely because sufficiently detailed data have not been available at that level. In this paper the development of stem-initial consonant alternation is studied with reference to the languages of Epi, in central Vanuatu, and the fusion hypothesis advanced by Lynch (1975) considered with respect to these languages. The final section of the paper examines the extent and range of stem-initial alternations within Vanuatu and their value as subgrouping evidence within this geographical area and within Oceanic as a whole.

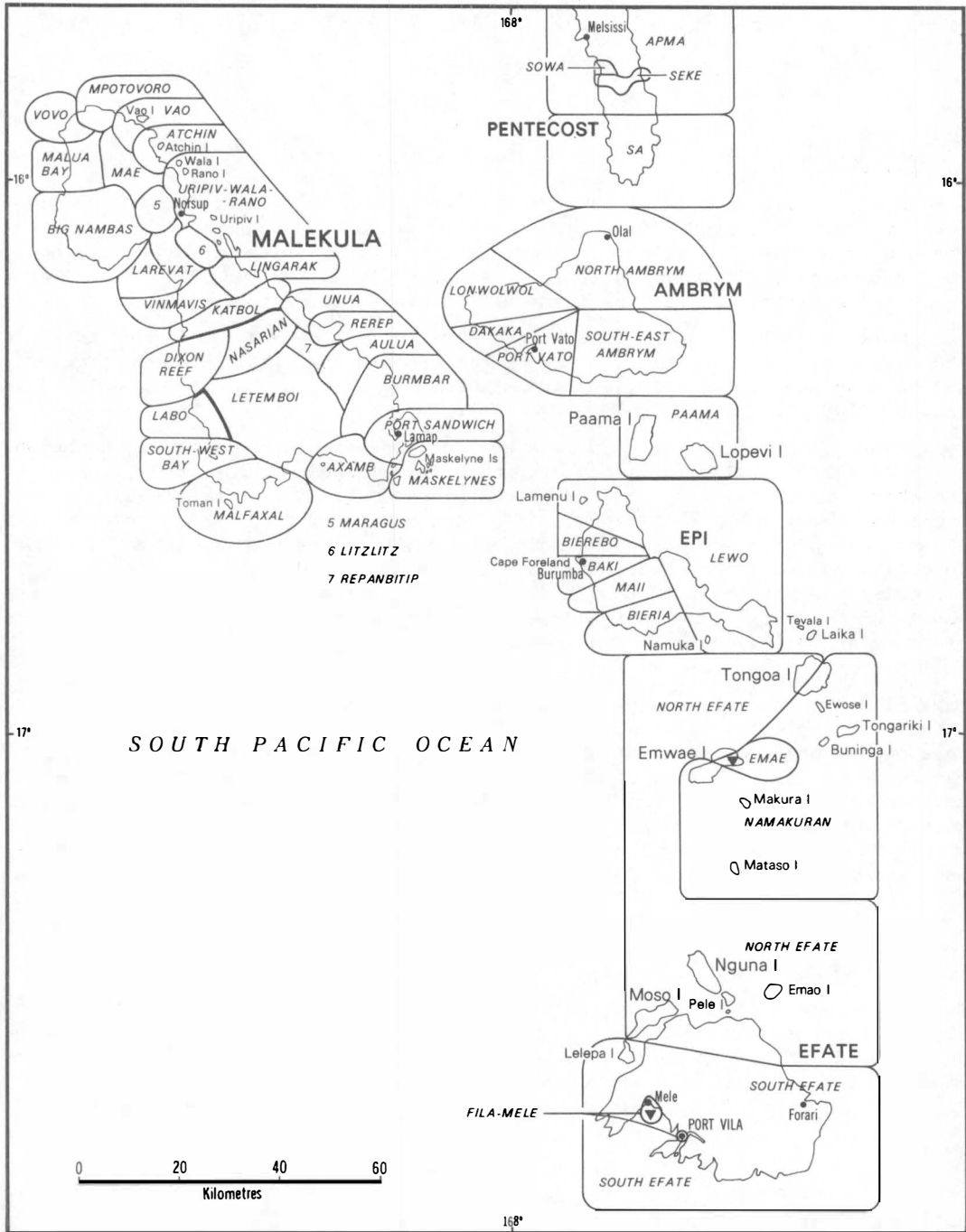
2. THE LANGUAGES OF EPI

There are currently five languages spoken on the island of Epi, namely Lewo, Bieria, Baki, Maii and Bierebo, as represented in Map 1. The languages and their principal dialects may be grouped according to the following family tree diagram, on both morphological and lexicostatistical evidence:

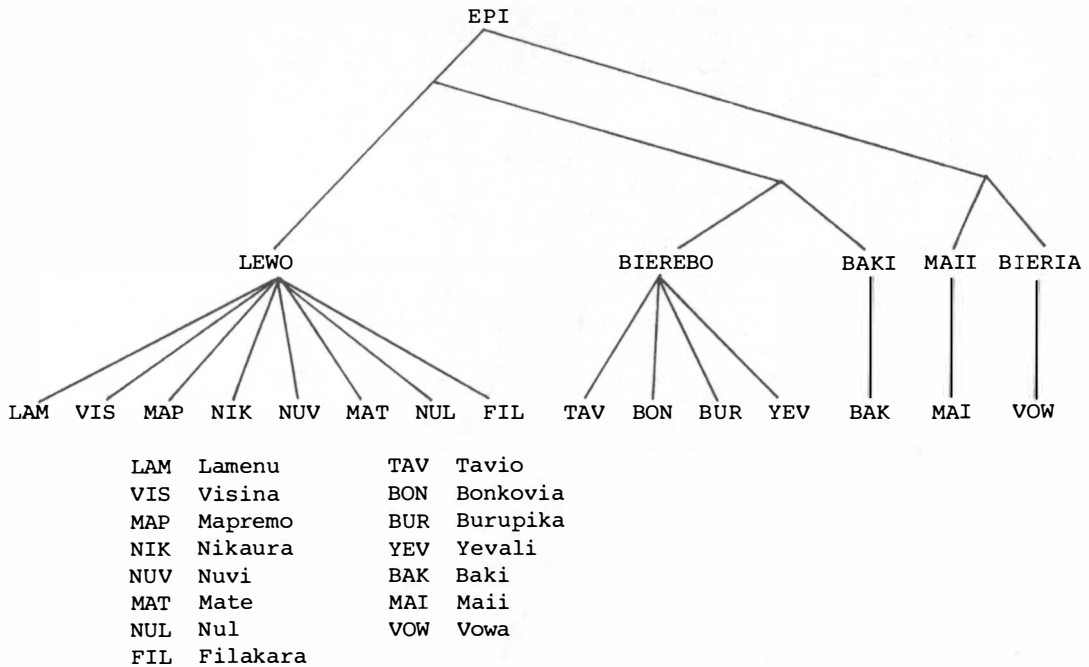
Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 239-258. *Pacific Linguistics*, C-94, 1986.

© D.T. Tryon

239

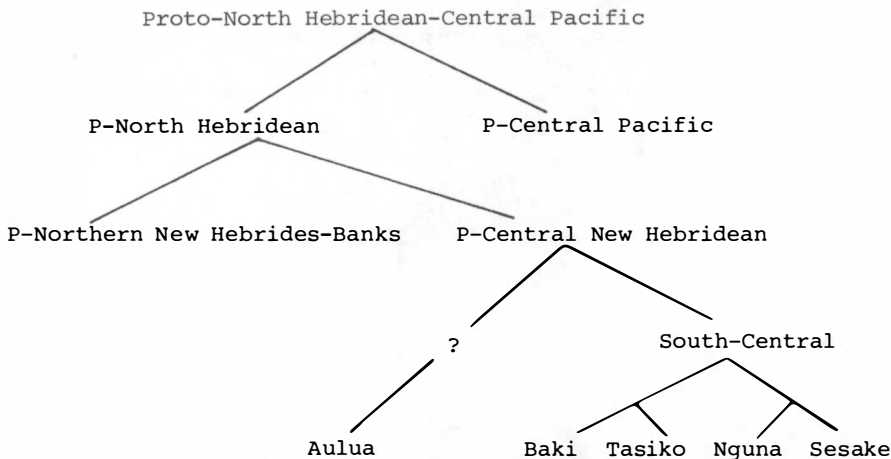


Map 1: Languages of Central Vanuatu



The family tree above does not represent the full extent of dialect variation within the languages of Epi, only the main dialects. Tables of sound correspondences are to be found in Table 2 (below).

In terms of their position within the Oceanic subgroup, Pawley (1972) considered that the Epi languages are members of a Central New Hebridean subgroup, which apparently also includes the languages of Malekula, as follows:



Pawley's classification, based on the tantalisingly little morphological data available at that time, has been developed by others more recently. The position of the Epi languages has remained relatively unaltered, however.

Tryon (1976), in a classification based largely on lexicostatistics, found that the languages of Epi form a single lower order subgroup within his East New Hebrides group, itself a subgroup of the huge North and Central New Hebrides group, the full extent of which his survey revealed for the first time.

Clark (1985) in a re-appraisal of the language situation in Vanuatu (formerly the New Hebrides), compared the subgroupings of Vanuatu languages put forward by Pawley and Tryon in their attempted classifications, re-examining the evidence. He concluded that all of the non-Polynesian languages of north and central Vanuatu constitute a single subgroup of Oceanic. With minor differences, then, Clark agreed with Pawley's interpretation of the available data. At the same time he found that the groupings indicated by the lexicostatistical evidence were largely supported by the morphological evidence which he adduced, with the exception of two lexicostatistically aberrant groups on Santo and Malakula, which are integrated in the North and Central Vanuatu group.¹

With respect to the Epi languages, Clark considers that they are most closely related to the languages of the Efate area, and that they constitute a primary branching within a Central Vanuatu subgroup, which subgroup includes the languages of Malakula, Pentecost, Ambrym, Paama, Epi and the Efate area.

While the Epi languages indeed share a few innovations with Efate, of a morphosyntactic nature, and while it appears likely that future research will confirm the existence of an Epi-Efate subgroup, the Epi languages share in two phonological developments which serve to separate them from the Efate languages, and indeed from most other languages in Clark's Central Vanuatu subgroup, namely the merger of POC *s and *ns as \emptyset in all of the Epi languages except Bieria where they merge as /h/, and the merger of POC *d and *l as /l/, whereas *d and *l do not merge in the Efate area.

3. THE ALTERNATIONS

The stem-initial consonant alternations in the languages of Epi are summarised in the following table:

Table 1: Summary table of consonant alternations

	LEWO	NUL	BIEREBO	BAKI	MAII	BIERIA
1.	p/v	p/ \emptyset	p/v	mb/v	mb/v	mb/v
2.	pw/w		pw/w			
3.			nd/t	c/t	md/t	nd/t
4.			nj/c	c/s		
5.			ŋg/k			
6.	miV/V	miV/V	miV/V	miV/V	miV/V	miV/V
7.	mC/C	mC/C	mC/C	mC/C	mC/C	mC/C
8.	C/C	C/C	C/C	C/C	C/C	C/C
9.						ŋ/h

It should be noted that Nul has been included with the other languages even though it is only a dialect of Lewo, as it contains one alternation which differs from Lewo, namely p/ \emptyset instead of p/v.

The left hand member of the alternation pair indicates the 'realis' in the languages listed, while the right hand member denotes an 'irrealis'. Examples

are given for each alternation in each language. Numbers beside the examples match the alternation number given in the summary table (Table 1). Examples:

(1) LEWO:

1. nakonevæ ne-pisu yaru nene
yesterday I-see(R) man that
I saw that man yesterday.

peni ne-visu yaru nene
tomorrow I-see(IRR) man that
I'll see that man tomorrow.

2. nakonevæ no-pwere kilika
yesterday I-pull(R) rope
I pulled the rope.

peni no-were kilika
tomorrow I-pull(IRR) rope
I'll pull the rope.

6. ne-m-iuveve
I-R-breathe
I breathed.

ne-uveve
I-breathe(IRR)
I shall breathe.

7. ne-m-loru
I-R-fall
I fell.

ne-loru
I-fall
I shall fall.

8. nakonevæ ni-mærau
yesterday I-fear
I was frightened yesterday.

peni ni-mærau
tomorrow I-fear
I'll be frightened.

(2) NUL:

1. konea ne-pen vila
yesterday I-go(R) Vila
I went to Vila yesterday.

pen n-en vila
tomorrow I-go(IRR) Vila
I'm going to Vila tomorrow.

6. ne-m-iu-ia kap
I-R-blow-O fire
I blew on the fire.

ne-iu-ia kap
I-blow(IRR)-O fire
I shall blow on the fire.

7. konea ne-m-leru
yesterday I-R-fall
I fell down yesterday.

pen ne-leru
tomorrow I-fall (IRR)
I shall fall tomorrow.

8. konea ne-mælio
yesterday I-sleep
I slept yesterday.

pen ne-mælio
tomorrow I-sleep
I shall sleep tomorrow.

(3) BIEREBO:

1. ne-ptali-a-kuñu
I-laugh(R)-O-me
I laughed at myself.

ne-vtali-a-kuñu
I-laugh(IRR)-O-me
I'll laugh at myself.

2. ɲeñowa na no-pwəl tebaka ia stoa
yesterday dem. I-buy(R) tobacco at store
I bought tobacco at the store yesterday.

pen na no-wəli-a
tomorrow dem. I-buy(IRR)-O
I'll buy some tomorrow.

3. nalə a-ndika namburu
they they-live(R) Namburu
They live at Namburu.

pen na ne-tika paama
tomorrow dem. I-live(IRR) Paama
I'll stay on Paama tomorrow.

4. ne-njoŋi-a
I-hear(R)-O
I hear it.

ne-coŋo lemawa
I-hear(IRR) try
I'll try to hear it.

5. kuli ɲgiriri
dog Ø-run(R)
The dog ran.

pen na kuli kiriri
tomorrow dem. dog Ø-run(IRR)
The dog will run.

6. ko-m-iliwi ku-mbano va-tano
you-R-stoop you-go(R) loc-down
You went underneath.

pen na ne-iliwi ne-vano va-tano
tomorrow dem. I-stoop(IRR) I-go(IRR) loc-down
I'll go underneath tomorrow.

7. ɲeñowa na ne-m-lae yumwa
yesterday dem. I-R-make house
I made a house.

pen na ne-lae yumwa
tomorrow dem. I-make house
I'll make a house.

8. ɲeñowa na ne-sei kærtiti
yesterday dem. I-wash baby
I washed the baby.

pen na ne-sei kærtiti
tomorrow dem. I-wash baby
I'll wash the baby.

(4) BAKI:

1. niovo nə-mbio-cau
yesterday I-call out(R)-you
I called to you yesterday.

pen nə-vio-cau
tomorrow I-call out(IRR)-you
I'll call to you tomorrow.

3. niovo ne-celan kia-mu tavak
yesterday I-light(R) Cl-your tobacco
I lit your cigarette.

pen ne-telan kia-mu tavak
tomorrow I-light(IRR) Cl-your tobacco
I'll light your cigarette.

4. ne-cen iup
I-eat(R) yam
I ate the yam.

ne-sen iup
I-eat(IRR) yam
I'll eat the yam.

5. nə-m-ial-cau
I-R-see-you
I saw you.

pen ne-al-cau
tomorrow I-see(IRR)-you
I'll see you tomorrow.

7. niovo nə-m-we kuli
yesterday I-R-hit dog
I hit the dog.

pen nə-we kuli
tomorrow I-hit dog
I'm going to hit the dog.

8. niovo nə-mro ka kuli
yesterday I-fear Inst dog
I feared the dog.

pen nə-mro ka kuli
tomorrow I-fear Inst dog
I'll fear the dog.

(5) MAII:

1. ləmbəŋə mbuar ilə-ŋ
dog Ø-bite(R) leg-my
The dog bit my leg.

tə ləmbəŋə vuar-tnau
fut. dog Ø-bite(IRR)-me
The dog will bite me.

3. nə-m-duluwa
I-R-call out
I called out.

tə nə-ruluwa
fut. I-call out
I shall call out.

6. nə-m-iop
I-R-see
I see/saw.

tə n-o·p
fut. I-see
I shall see.

7. inəvəmə nə-m-ləŋə n-tai m-dum
yesterday I-R-hear art-sea Ø-R-roar
I heard the sea roar yesterday.

tə nə-ləŋə ra·mbiə
fut. I-hear tomorrow
I'll hear it tomorrow.

8. inəvəmə nə-ma n-top
yesterday I-eat art-sugarcane
I ate sugarcane yesterday.

tə nə-ma n-top ra·mbiə
fut. I-eat art.sugarcane tomorrow
I'll eat sugarcane tomorrow.

(6) BIERIA:

1. nove no-mbek va
yesterday I-go(R) home
I went home yesterday.

tambia ne-vek va
tomorrow I-go(IRR) home
I'm going home tomorrow.

3. nove no-ndokosan
yesterday I-sit down(R)
I sat down yesterday.
- tambia ne-tokosan
tomorrow I-sit down(IRR)
I shall sit down tomorrow.
6. nove no-m-iat mbukah
yesterday I-R-tie pig
I tied up the pig yesterday.
- tambia ne-at mbukah
tomorrow I-tie pig
I'll tie up the pig tomorrow.
7. nove no-m-te lakai
yesterday I-R-cut wood
I cut the wood yesterday.
- tambia ne-te lakai
tomorrow I-cut wood
I'll cut the wood tomorrow.
8. nove no-matak
yesterday I-fear
I was frightened yesterday.
- tambia ne-matak
tomorrow I-fear
I shall be frightened tomorrow.
9. nove no-m-ul le-ŋ
yesterday I-R-rub leg-my
I rubbed my leg yesterday.
- tambia ne-hul le-ŋ
tomorrow I-rub leg-my
I'll rub my leg tomorrow.

Before moving on to discuss the development of the alternation system within the languages of Epi, a few comments on the examples presented above are given which are relevant to this discussion.

In all of the Epi languages there is a realis marker of the form *mV which occurs, potentially, between the subject marker and the verb stem. Thus we have a common verb phrase structure for realis as follows:

SM + mV + verb Stem † Trans.Suffixes etc.

However, the realis marker is only used with a certain number of verb-stem initial consonants. In other cases, there is no overt reflex of *mV, but rather an alternation, usually of an oral/nasal grade nature, between stem-initial consonants. This process has been fully illustrated in the examples given above, from all of the languages. The distribution of the alternations and/or overt realis marker will be discussed below. Suffice it to say at this point that the Epi realis/irrealis marking system is quite regular.

It should be noted that Bieria has developed a double indication of realis/irrealis in that the vowel of the subject marker differs from realis to irrealis. Thus:

no-matak *I am afraid*
 ne-matak *I shall be afraid*

This distinction is maintained in Bieria also in cases where there is a verb-stem initial consonant alternation, as the examples above show. This feature has also been noted with some verb stems in Baki. An explanation will be offered in the next section, on the historical development of the alternations.

The consonantal alternations are not restricted to verb stems, but also apply in other environments within the verb phrase. In Maii, for example, they appear with the multiplicative mbaka-/vaka-, and with the inceptive mba·lik/va·lik. Examples:

funggo² mbaka-tol
 ø-do mult-three
He did it three times.

tə funggo vaka-tol
 fut. ø-do mult-three
He will do it three times.

nə-frei əvrə ndə mba·lik
I-write thing only start(R)
I began to write things.

mə-mbənəmə iə ndə ukau ŋei mba·lik
we exc-come(R) in only year this start(R)
We came and started only this year.

nə-fluŋ ndə nə-vənəmə vila va·lik nə
I-want only I-come(IRR) Vila start(IRR) this
I just want to come to Vila for the first time.

Consonantal alternations also occur with the morphemes indicating indirect object and accompaniment, as in the following:

m-kun mbikin-tnau
 ø-R-give to-me(R)
He gave it to me.

tə kun vikin-tnau
 fut. ø-give to-me(IRR)
He will give it to me.

nə-m-i·mbə mblaken a-ŋ ləmbaŋə
I-R-go with(R) Cl-my dog
I went with my dog.

tə n-i·və vlaken a-ŋ ləmbaŋə
 fut. I-go with(IRR) Cl-my dog
I shall go with my dog.

The base form of the consonantal alternations in the languages of Epi is the irrealis, for this is the unmarked form. It is also from this form that nouns are formed from verb bases, for example, Maii:

mbivi/vivi to work vivi-yanə work
 mbelau/velau proud velo-wanə pride

(Note: the /y/ form of the noun formative is used following non-back vowels, while the /w/ form is used elsewhere.)

With negatives, the irrealis form of the verb stem is used, as in the following examples from Maii:

nə-vuk voi suksuk
I-read(IRR) not book
I did not read the book.

nə-ta voi lakai
I-cut(IRR) not wood
I did not cut the wood.

4. DEVELOPMENT OF STEM-INITIAL CONSONANT ALTERNATIONS

Lynch (1975) proposed a series of five steps to account for the development of stem-initial consonant alternations in Nguna (Vanuatu) as follows:

1. PHC *ma is a free preverbal particle marking realis:
 Thus: Pre-Nguna *ma doŋo (realis *hear*)
2. PHC *ma becomes a clitic (normally prefixed to the verb):
 Thus: Pre-Nguna *ma-doŋo (realis *hear*)
3. The vowel of the prefix is lost:
 Thus: Pre-Nguna *m-doŋo (realis *hear*)
4. The /m/ assimilates to the following consonant:
 Thus: Pre-Nguna *n-doŋo (realis *hear*)
5. Prenasalised consonants may undergo subsequent sound changes:
 Thus: Nguna toŋo (realis:*hear*)

Lynch thus explains the development of the alternating Nguna pair roŋo/toŋo as follows:

irrealis	realis	
*doŋo	*m(V)-doŋo	
*doŋo	*n-doŋo	(assimilation)
roŋo	toŋo	(change *d > r *nd > t)

The above analysis of a language with really very little consonant alternation, and based on not a great deal of data, suggested that the languages of Epi, where stem-initial consonant alternation plays a major part, should be investigated with respect to this feature. Examples from all of the Epi languages have been presented above. In order to trace the development of the alternations within Epi, a large number of verb stems from each of the languages was examined in terms of their derivation from proto-Epi, the sound correspondences for which are given in Table 2. The derived pairs are set out in Table 3 below. Some reconstructed forms are given in Table 4.

A comparison of the sound correspondences and the alternating pairs of verb stems in Table 3 shows that the irrealis stem forms (the left hand member of each pair) are regular reflexes of the proto-forms, while the realis verb stem forms do not follow the regular sound correspondences. Indeed, the phonemic status of the voiced members of the consonant pairs in the languages of Epi may be questioned according to some phonological theories since it is extremely difficult to discover voiced/voiceless stop oppositions which do not involve realis/irrealis verb stem oppositions.

Table 2: Epi sound correspondences

POC	*p	*mp	*t	*nt	*k	*ŋk	*d	*nd	*q	*ŋp
PEP	*p	*b	*t	[*d]	*k	*g	*l	*l	∅	*pw
LAM	v,p		t,s	[s]	k		l	l	∅	p
VIS	v,p	p	t	t	k		l	l	∅	p
MAP	v,p	b	t	t	k	k	l	l	∅	p
NIK	v,p	p	t	t	k	g	l	l	∅	p
NUV	v,p	p	t	t	k		l	l	∅	p
MAT	∅,p	p	t	t	k		l	l	∅	pw
NUL	∅,p	p	t	t	k		l	l	∅	pw
FIL	∅,p	p	t	t	k		l	l	∅	pw
TAV	v,p	p	t	t	s,k ¹	g	l	l	∅	v
BON	v,p	p	t	t	s,k ¹	g	l,c ²	c ³	∅	v
BUR	v,p	b	t	t	s,k ¹	g	l,c ²	c ³	∅	p
YEV	v,p	b	t	t	k	g	l,c ²	c ³	∅	p
BAK	v,p	b	t	t	s,k ¹	k	l,c ²	l	∅	bw
MAI	v,p	b	t	[d]	k	k,ŋ	l	l	∅	p
VOW	v,p	b	t,s	[r]	k	k,ŋ	l	l	∅	b

POC	*s	*ns	*m	*n	*ñ	*ŋ	*l	*R	*w	*y
PEP	*h	*h	*m	*n	*ñ	*ŋ	*l	*R	*w	∅
LAM	∅	∅	m	n	n	ŋ	l	l,∅	w	
VIS	∅	∅	m	n	n	ŋ	l	l,∅	w	
MAP	∅	∅	m	n	n	ŋ	l	l,∅	w	
NIK	∅	∅	m	n	n	ŋ	l	l,∅	w	
NUV	∅	∅	m	n	n	ŋ	l	l,∅	w	
MAT	∅	∅	m	n	n	ŋ	l	l,∅	w	
NUL	∅	∅	m	n	n	ŋ	l	l,∅	w	
FIL	∅	∅	m	n	n	ŋ	l	l,∅	w	
TAV	∅	∅	m	n	n	ŋ	l	l,∅	w	
BON	∅	∅	m	n	ñ	ŋ	l,c ²	l,∅	w	∅
BUR	∅	∅	m	n	ñ	ŋ	l,c ²	l,∅	w	
YEV	∅	∅	m	n	ñ	ŋ	l,c ²	l,∅	w	
BAK	∅	∅	m	n	ñ	ŋ	l,c ²	l,∅ ⁴	w	
MAI	∅	∅	m	n	n	ŋ	l	∅	w	∅
VOW	h	h	m	n	n	ŋ	l,∅	∅	w	

¹With pronouns.
²[c] initial; [l] intervocalic.
³Examples reflect *nd in word-initial position only.
⁴POC *R reflected as /l/ throughout Epi only reflecting a few POC items; otherwise *R is reflected as zero.

Table 3: Epi stem-initial consonant alternations

	LEWO	NUL	BIEREBO	BAKI	MAII	BIERIA
1. *p/*m-p	<u>v/p</u> vano/pano go	<u>ɔ̃/p</u> en/pen go	<u>v/p</u> vtali/ptali laugh	<u>v/mb</u> vio/mbio call out	<u>v/mb</u> vuar/mbuar bite	<u>v/mb</u> vek/mbek go
2. *pu/*m-pu	<u>w/pw</u> were/pwere pull		<u>w/pw</u> weli/pweli buy			
3. *t/*m-t	<u>s/s</u> [*tu] sulua/sulua sew <u>t/t</u> taji/taji cry	<u>s/s</u> [*tu] smalu/smalu stand <u>t/t</u> te/te cut	<u>t/nd</u> teŋi/ndenji cry tika/ndika stay	<u>t/č</u> teŋi/čenji cry telan/čelan light	<u>t/md</u> ta/m-da cut <u>t/mt</u> tul/m-tul sew	<u>t/nd</u> tokosan/ndokosan sit <u>t/mt</u> te/m-te cut <u>s/ms</u> [*tu] sul/m-sul sew
4. *d/*m-d	<u>l/ml</u> lonɛa/m-lonɛa hear	<u>l/ml</u> lonɛa/m-lonɛa hear	<u>č/nj</u> čoŋo/njoŋo hear	<u>č/mč</u> čoŋ/m-čoŋ hear	<u>l/ml</u> lonɛa/m-lonɛa hear	<u>l/ml</u> lonɔ/m-lonɔ hear
5. *k/*m-k	<u>k/k</u> karia/karia bite kilia/kilia know	<u>k/k</u> kania/kania eat kilia/kilia know	<u>s/s</u> sani/sani eat sari/sari bite	<u>s/č</u> sen/čen eat	<u>k/mk</u> kan/m-kan eat	<u>k/mk</u> kan/m-kan eat
6. *g/*m-g			<u>k/ng</u> kiriri/ngiriri run			
7. *s/*m-s						<u>h/m</u> hul/mul rub
8. *V/*mi-V	<u>V/miV</u> uveve/mi-uveve breathe	<u>V/miV</u> iu/mi-iu blow	<u>V/miV</u> iliwi/mi-iliwi stoop	<u>V/miV</u> al/mi-al see	<u>V/miV</u> op/mi-op see	<u>V/miV</u> at-mi-at tie
9. *l/*m-l	<u>l/ml</u> lilua/m-lilua vomit	<u>l/ml</u> lua lua/m-lua lua vomit	<u>č/nj</u> čulua/njulua vomit	<u>č/mč</u> čuluo/m-čuluo vomit	<u>l/ml</u> lua lua/m-lua lua vomit	<u>l/ml</u> luawa/m-luawa vomit
10. *w/*m-w	<u>w/mw</u> we/m-we hit	<u>w/mw</u> we/m-we hit	<u>w/mw</u> we/m-we hit	<u>w/mw</u> we/m-we hit		
11. *m/*m-m	<u>m/m</u> mai/mai be sick	<u>m/m</u> munia/munia drink	<u>m/m</u> mrau/mrau fear	<u>m/m</u> mun/mun drink	<u>m/m</u> mar/mar die	<u>m/m</u> mat/mat die

Table 4: Some proto-Epi verb root reconstructions

POC	PEP	Bierebo	Baki	Lewo	Nul	Maii	Bieria
*-	*te [cut]	nde/te	ce/te	te/te	te/te	da/ra	te/te
*toka	*toko [be]	ndo/to	co/to	to/to	te/te	də/rə	doko/toko
*taŋis	*taŋi [cry]	ndeŋi/ teŋi	ceŋi/ teŋi	taŋi/ taŋi	taŋi/ taŋi		
*pano	*vano [go]	mbano/ vano	mbano/ vano	pano/ vano	pen/en	mbəne/ vəne	mbek/vek
*puti	*vuto [pull]	pur/wur	mburo/ vuro	pure/ vure	pur/ur		
*-	*ve [be]	pe/ve	mbe/ve	pe/ve	pe/e	mbe/ve	
*-	*vuyu [climb]		mboyo/ voyo	puyu/ vuyu	pu/u		
*poli	*vuli [buy]	puli/ wuli	mbulio/ vulio	pulia/ vulia		mbul/vul	
*putu	*vut(o,e) [flee]	mbure/ wure	mburo/ vuro	pure/ vure			
*kani	*kani [eat]	sanial/ sanial	cen/sen	kaniak/ kaniak	kaniak/ kaniak	kan/kan	kan/kan
*kati	*kati [bite]	sarial/ sarial	cer/ser	kariak/ kariak	kariak/ kariak		
*-	*kuḡari [scratch]	saral/ saral	cubwar/ cubwar	kupariak/ kupariak	kupariak/ kupariak		
*-	*kove [throw]	soveal/ soveal		koveniak/ koveniak	koveniak/ koveniak	kəvək/ kəvək	kovək/ kovək
*doŋo	*loŋo [hear]	njoŋ/ coŋ	coŋ/coŋ	loŋeak/ loŋeak	loŋoia/ loŋoia	ləŋə/ləŋə	loŋo/loŋo
*luaq	*lualua [vomit]	njulua/ culua	culuo/ culuo	lilua/ lilua	lualua/ lualua	lualuə/ lualuə	luawa/ luawa
*-	*loru [fall]	njoru/ coru	coru/ coru	loru/ loru	leru/ leru		
*-	*kitikit [run]	ŋgiriri/ kiriri	cikiti/ tikiti	kiriri/ kiriri	kiriri/ kiriri	kirkir/ kirkir	kitikit/ kitikit

The course of development proposed by Lynch for Nguna appears to be substantiated in the case of the languages of Epi. If, for example, one were to consider the development of the consonantal pairs set out in Table 3, one would recognise that they follow a regular developmental sequence, as follows:

- (a) With all Epi languages a form *mV preceded all verb stems to indicate realis, becoming prefixed throughout.
- (b) In the case of a number of stem-initial consonants, especially those reflecting proto-Epi *p, *t, *d and to a lesser extent *k and *g, stem-initial alternations developed as a result of the interaction of the realis marker and the initial consonant of the verb stem.
- (c) With other stem-initial consonants, especially those reflecting proto-Epi *l, *w and *m, the presence of the realis marker produced only the sequences m-l, m-w and m-m.

The expected phonological developments detailed in the Lynch-proposed developmental sequence will be seen to have taken place in the Epi languages, individually at least, for example:

*m+p > mv > mb > mp > p (Lewo: vano/pano *go*)
 *m+t > md > nd > nt > t (Bierebo: teŋi/ndeŋi *ery*)³

While the current Epi languages may be presumed to have followed the path of assimilation of the initial /m/ of the realis marker *mV to the initial consonant of the verb stem, once the vowel of the realis marker was lost (which would be normal in unstressed position, no single set of developmental phonological rules has been worked out which covers all of the languages. This suggests that the stem-initial consonant alternations in Epi are likely to have been independent phonological developments within the island. As such, then, the alternations per se may not be of very great value as subgrouping evidence. On the other hand, the fact that it is the irrealis form which is unmarked and which is demonstrably the base form may well prove a useful piece of evidence in this regard, especially when languages outside the Epi-Efate subgroup are considered.

It will have been observed, also, that while the development of the stem-initial consonant alternations has proceeded regularly within the Epi languages, there are irregularities within both Maii and Bieria with respect to the development of *m+t, as follows:

- (a) In Bieria, incidentally the most closely related language to Maii, as well as the t/nd stem-initial alternation, from *t/*m-t, we also have a few t-initial stems where the collocation of *m+t has not produced a consonantal alternation. Rather the assimilatory chain has not developed past stage one. Thus we have, for example:

	ne-tokosan	<i>I shall sit down</i>
	no-ndokosan	<i>I sat down</i>
but:	ne-te	<i>I shall cut</i>
	no-m-te	<i>I cut</i>

As noted above, Bieria doubly marks realis/irrealis with many verbs first by the stem-initial consonant alternation but also by the vowel of the subject marker. This latter phenomenon may well have developed as a device to distinguish tense/aspect with verb stems whose initial consonants do not participate in the alternation process, for example those beginning with m-, l- or w-.

- (b) In Maii, Bieria's neighbour, we have the same situation, only one stage earlier than in Bieria. For in Maii we have the same lack of development with respect to some t-initial verb stems, namely *m+t remains unaltered. Thus:

ne-tul *I shall sew*
 ne-m-tul *I sewed*

At the same time, many t-initial verb stems have taken the next step in the developmental process, namely *m-t becomes Maii /md/, for example:

nə-ta· lakai
I-cut(IRR) wood
I shall cut the wood.

nə-m-da· lakai
I-R-cut wood
I cut the wood.

The picture in Maii and Bieria, then, suggests that in respect of those two languages at least the developmental process has been caught in midstream and that it is therefore still an active process and that the phenomenon of stem-initial consonant alternation in the languages of Epi is a relatively recent development.

5. STEM-INITIAL ALTERNATIONS OUTSIDE EPI

Stem-initial consonant alternations occur in a number of languages outside Epi. They are quite rare outside Vanuatu, having been reported only for Yabem and Mapos Buang in Papua New Guinea among the Oceanic subgroup of Austronesian (Lynch 1975:91). They are not of direct concern here as the alternations in those languages result from the assimilation of an irrealis marker *na with initial consonants of the verb stems. Within Vanuatu, they all appear to have originated as a result of the assimilation of a realis marker *mV to the initial consonant of some verb stems.

Within Vanuatu stem-initial consonant alternations are encountered in Erromango, and in a geographical continuum from Efate to Pentecost and Ambae (Aoba).

- In Efate one has, for example:

e pano *he goes*
 e pe vano *if he goes* (Schütz 1969)

- In Namakura (in the Efate area) one has, for example:

ni-mboh naim *I build a house*
 ni-mba-woh naim *I shall build a house*
 ni-ndah *I come back*
 ni-mba-rah *I shall come back* (own data)

However, in Namakura one has also verb phrases such as:

ni-mba-maha-mboh
I-fut.-again-do(R)
I shall do it again.

ni-mba-maha-ndah
I-fut.-again-come back(R)
I'll come back again.

- To the north of the Epi-Efate subgroup, one finds in Paama, for example:

a-daasili
they-disperse(R)
They dispersed.

i-taasili
they-disperse(IRR)
They will disperse. (Crowley 1982)

In Paama, however, Crowley notes that there are irregularities in the system, whereby the same initial consonant produces different alternations. He cites, for example:

titilu Class 1 *sew* (d- (R)/t- (IRR))
 titilu Class 5 *drip* (t- (R)/t- (IRR)) (1982:121)

It may be, however, that the initial /t/ in the two forms derives from two different proto-phonemes.

- In the languages of North Ambrym, stem-initial alternations are also present, as in:

na-m-tivi
I-R-cut
I have cut.

na-te-rvi
I-past-cut
I cut (past).

na-ŋa-rvi
I-fut.-cut
I shall cut. (own data)

- On Pentecost, alternations are found throughout the languages. In Apma, for example, Walsh (1982:238) reports:

na-m-beb
I-act. in prog.-say
I say

ra-t-veb
they-comp.act.-say
They said

na-m-ban
I-act. in prog.-go
I go

na-t-van
I-comp.act.-go
I went

In Raga, further to the north, Walsh provides very full evidence of a number of stem-initial consonant alternations, for example:

na-m-bano
I-act. in prog.-go
I go

ra-m-bano
they-act. in prog.-go
They go

na-n-van-vano
I-cont.-go-go
I used to keep on going. (Walsh 1982:237)

- Alternations are also reported on Ambae (Aoba) by Ivens, for the Lobaha dialect (N.E. Ambae), citing forms such as:

vava/bava to speak
vora/bora to become (Ivens 1939-42:356)

- In southern Vanuatu, outside the Northern-Central Vanuatu language subgroup, Sie, on Erromanga, also exhibits stem-initial vowel and consonant alternation. Lynch (1975:91) records, for example:

past	non-past	
oyhi	aŋhi	to see
okili	aŋkili	to know
oruh	aduh	to swim
ovelam	ampelam	to come
etur	antur	to stand

Within the Northern-Central Vanuatu subgroup it is significant that the verb base from which nouns are formed reflects the oral grade consonant throughout. Examples:

Nguna:	na-vasa-ana	speaking	(vasa/pasa)
Namakura:	tokean	village	(tok/ndok) [stand]
	na-worean	story	(wor/mbor)
Maii:	vivi-ane	work	(vivi/mbivi)
Paama:	hilu-ene	cough	(hilu/vilu)
Raga:	noyu vano-ana	my going	(vano/pano)

6. CONCLUSION

A number of conclusions may be drawn from the data presented above, the principal ones being:

(1) Within Vanuatu a distinctive verb-stem initial consonant alternation occurs, within the Northern-Central Vanuatu subgroup area, in a geographical continuum stretching from Efate in the south to Ambae in the north. There is also a suggestion (Pawley 1972, Clark 1985) that the alternations may have once been present or perhaps remain to be discovered in the almost undescribed languages of Malakula, where traces appear in the numerals and with certain adjectives.

(2) Vowel and stem-initial consonant alternations also occur in Sie (Erromanga). Sie is outside the Northern-Central Vanuatu subgroup, being part of the Southern Vanuatu subgroup (Lynch 1978). The Sie alternations appear to be unrelated to and rather different from those observed in some Northern-Central Vanuatu languages, as do those reported for Yabem and Mapos Buang in Papua New Guinea.

(3) The alternations arise from the presence of a 'realis' marker *mV preceding the verb stem. The data presented from the Epi languages shows that there is no single integrated developmental process from proto-Epi to the present day languages, suggesting strongly that the consonant alternations there and by implication elsewhere in Vanuatu are the result of independent parallel development.

(4) Evidence from outside Epi, particularly that from Pentecost (Raga and Apma) and Ambrym, shows that the alternations have a phonological rather than a morphological basis, for wherever one finds a reflex of *mV preceding a verb stem (given that alternations occur only with certain consonants in any given language) an alternation follows, despite the fact that other realis/past markers co-occur in the same language.

(5) Despite the fact that the stem-initial consonant alternations are probably the result of independent parallel development, the fact that the base or unmarked form, and the form from which nouns are derived from verbs, is the irrealis throughout the languages manifesting consonant alternation is rather remarkable and one which will prove valuable as the languages of the Northern-Central Vanuatu subgroup become better known. For while the alternations are now known to extend beyond the Central Vanuatu subgroup (Clark 1985) into the Northern Vanuatu subgroup, the possibility still exists that the phenomenon may be discovered beyond the Northern and Central Vanuatu subgroup and may be assigned to some higher level subgrouping within Oceanic. Indeed, Lynch and Tryon (1985) have shown that reflexes of the 'realis' marker *mV extend beyond Vanuatu, increasing the possibility of further discoveries in terms of form and/or function of stem-initial consonant alternations within Oceanic.

NOTES

1. The quantitatively based groupings fit remarkably well with those based on qualitative evidence, with the two minor exceptions mentioned. Since morphological evidence of the type adduced by Clark, and before him Pawley, is more weighty than lexicostatistical evidence, and since that evidence points to the inclusion of the Interior Malekula and East Santo in the Northern-Central subgroup, it is preferable to include both of those lexicostatistically indicated subgroups within Northern-Central Vanuatu.
2. f-initial stems do not undergo alternation in Maii.
3. The Bierebo forms exemplify an interstage development; others may be found in the examples set out in Table 4.

REFERENCES

- CLARK, R.
 1985 Languages of north and central Vanuatu: groups, chains, clusters and waves. In Andrew Pawley and Lois Carrington, eds *Austronesian linguistics at the 15th Pacific Science Congress*, 199-236. PL, C-88.
- CODRINGTON, R.H.
 1885 *The Melanesian languages*. Oxford: Clarendon.
- CROWLEY, T.
 1982 *The Paamese language of Vanuatu*. PL, B-87.
- IVENS, W.G.
 1939 A grammar of the language of Lamalanga, North Raga, New Hebrides. *BSOAS* 9/3:733-763.
 1942 A grammar of the language of Lobaha, Lepers' Island, New Hebrides. *BSOAS* 10/1:345-363.
- LYNCH, J.D.
 1975 Oral/nasal alternation and the realis/irrealis distinction in Oceanic languages. *Oceanic Linguistics* 14/2:87-99.
- LYNCH, John and D.T. TRYON
 1985 Central-Eastern Oceanic: a subgrouping hypothesis. In Andrew Pawley and Lois Carrington, eds *Austronesian linguistics at the 15th Pacific Science Congress*, 31-52. PL, C-88.
- PAWLEY, A.K.
 1972 On the internal relationships of the Eastern Oceanic languages. In R.C. Green and M. Kelly, eds *Studies in Oceanic culture history*, vol.3, 1-142. Pacific Anthropological Records No.13. Honolulu: Bernice P. Bishop Museum.
- RAY, S.H.
 1926 *A comparative study of the Melanesian island languages*. Cambridge: Cambridge University Press.
- TRYON, D.T.
 1973 Linguistic subgrouping in the New Hebrides: a preliminary approach. *Oceanic Linguistics* 12:303-351.
 1976 *New Hebrides languages: an internal classification*. PL, C-50.
- WALSH, D.S.
 1982 Variation of verb-initial consonants in some Eastern Oceanic languages. In Amran Halim, Lois Carrington and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, vol.1: *Currents in Oceanic*, 231-242. PL, C-74.

THE PROTO-SOUTHERN VANUATU PRONOMINAL SYSTEM

John Lynch

0. INTRODUCTION

This paper explores the development of the Proto-Oceanic (POC) pronominal system and pronominal forms in languages of the Southern Vanuatu (SV) subgroup. It will focus particularly on the development of the forms themselves, and on the nature of the reconstructed system.

0.1 The Southern Vanuatu languages

The SV languages form a closed subgroup of Oceanic, composed of the non-Polynesian languages of the Tafea District of Vanuatu;¹ (West) Futuna-Aniwa, a Polynesian Outlier, is the only language of the District which is not a member of the subgroup. The languages (with three-letter abbreviations), their locations by island, the approximate number of speakers, and major sources of data for each language are shown in Table 1. Additional geographical information is given on the accompanying map.

Table 1: Languages of the Southern Vanuatu subgroup

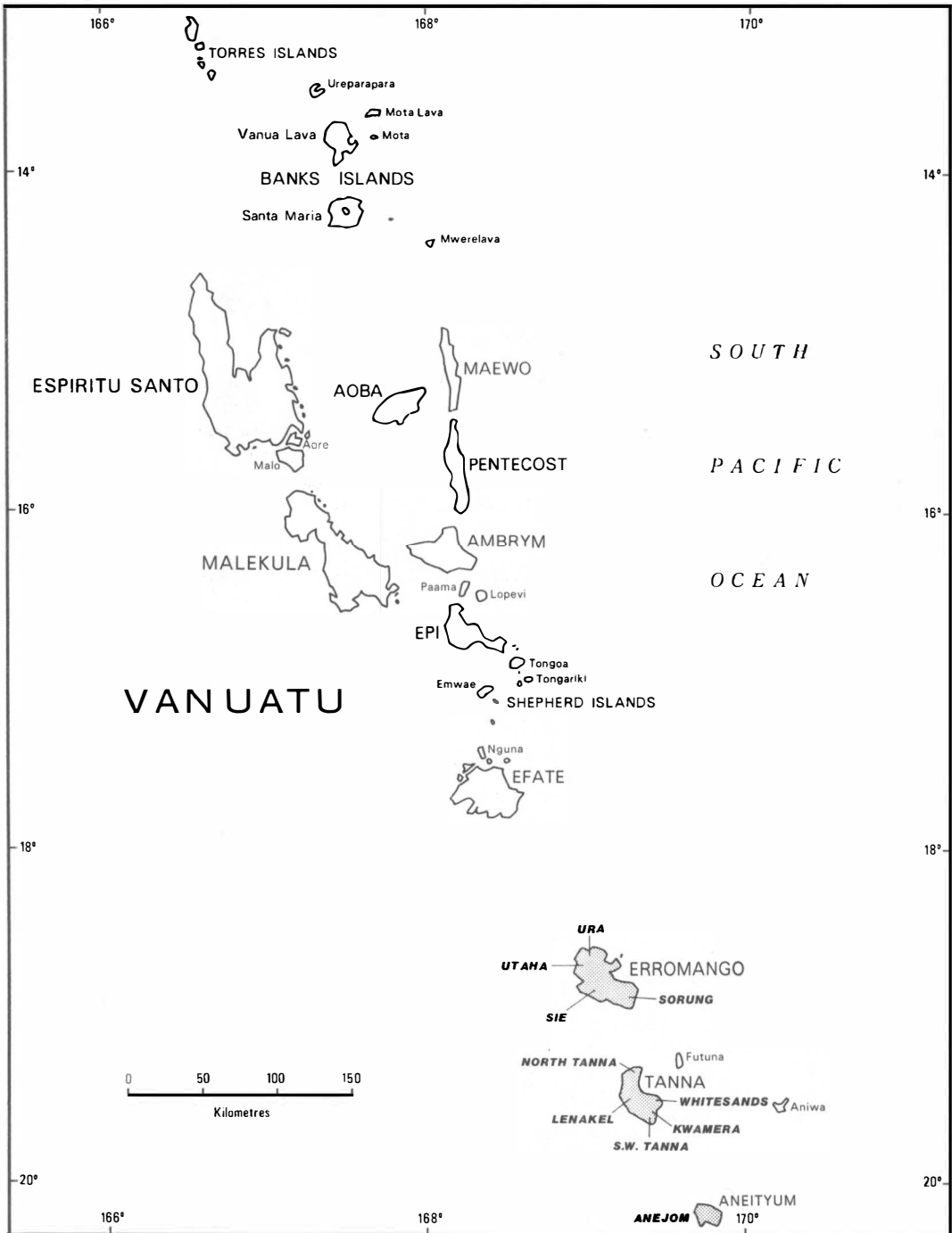
Island	Language	Abbrev.	Popul'n	Major data sources
Erromango	Sie	SIE	900	Lynch and Capell 1983, Capell and Lynch 1983
	Ura	URA	10	Lynch 1983b
	Utaha	UTH	extinct	Lynch 1983c
	Sorong	SOR	extinct	Lynch 1983d
Tanna	North Tanna	NTN	3,000	Lynch 1974, fieldnotes
	Whitesands	WSN	3,000	Lynch 1974, fieldnotes
	Lenakel	LEN	4,000	Lynch 1975,1977a,1978b
	South-west Tanna	SWT	2,000	Lynch 1982a
Aneityum	Kwamera	KWM	2,000	Lindstrom 1984
	Anejom	ANJ	500	Lynch 1982a

An examination of the internal relationships of these languages, focussing particularly on the development of POC phonology, was attempted in Lynch 1978a. That paper established the SV group as a whole, and also showed that three subgroups could be identified, one per island. Further, no evidence could be found

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 259-287. *Pacific Linguistics*, C-94, 1986.

© John Lynch

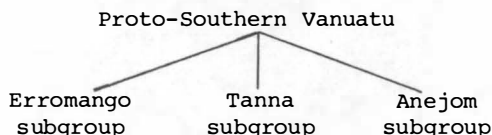
259



Map 1: Vanuatu

to combine any two of the subgroups together as against a third; that is, the data then available implied the family-tree as shown in (1):

(1)



More recent work on the interrelationships of the (extant and extinct) Erromangan languages (Lynch, ed. 1983, Lynch 1983e) in no way contradicts this picture.

The position of the SV group within the Oceanic family is not clear. Recent work (Lynch and Tryon 1983) suggests that the SV languages group with Pawley's (1972) Eastern Oceanic group, and possibly also with the languages of Utupua and Vanikoro, the Loyalty Islands and New Caledonia, in a 'Central Oceanic' group. While the status of this grouping is far from certain, an examination of recent work by Clark (1983, n.d.) on the languages of North and Central Vanuatu has convinced me that the SV languages are more closely related to their northern neighbours than I had originally thought.

0.2 Orthography, abbreviations and conventions

In general, I follow the orthography of the published sources listed in Table 1. For the sake of consistency or simplicity, however, some minor changes to those orthographic conventions have been made in this paper. They are:

1. pw, mw for Anejom b, \tilde{m} (velarised bilabials);
2. θ for Anejom d (voiced dental fricative);
3. γ for Anejom c and for all Erromangan languages g (voiced velar fricative);
4. η for Anejom and all Tanna languages g (velar nasal);
5. ə for all Tanna languages ɨ (non-low central vowel).

Note also that, in citing Erromangan forms, b, d and g represent prenasalised voiced stops.

In glossing pronominal forms, I borrow a number of conventions from Geraghty (1983:6-7). Pronominal forms are glossed with the following abbreviations:

I	first person	2	dual
Ii	first person inclusive	3	trial
Ix	first person exclusive	p	plural
II	second person	ns	non-singular (unspecified as to dual/trial/plural)
III	third person		

Note that a gloss like iik 'II' implies second person *singular*; non-singular numbers are always overtly glossed - e.g. kamilau 'II2'.

Other abbreviations used in glossing examples in the text are:

AOR	aorist	POSS	possessive marker
CONC	concurrent tense	PRES	present
CONT	continuative	PST	past
INCH	inchoative	SM	subject marker
IRR	irrealis	TR	transitive
MP	mid past		

1. FORM AND FUNCTION

A full list of all pronominal forms is given in the tables in section 2. Based on these data, this section will briefly examine the person, number and case distinctions found in SV pronominal forms, and will also comment briefly on the internal morphological structure of those forms.

1.1 Person

All SV languages distinguish three persons in the singular and four in the non-singular pronouns, the common Austronesian inclusive/exclusive distinction being found throughout the group. The following sets of focal pronouns illustrate this:

(2)	SIE		ANEJOM	
		koh Iip		akajau Ii2
	yau I	kam Ixp	añak I	ajamrau Ix2
	kik II	kimi IIp	æk II	ajourau II2
	iyi III	iror IIIp	aen III	aarau III2

1.2 Number

Anejom and all the Tanna languages distinguish four numbers in focal, possessive and (where they occur) objective pronouns: singular, dual, trial (marking three only) and plural. This is illustrated below:

(3) LENAHEL (focal)	S.W. TANNA (objective)	ANEJOM (possessive)
io I	in III	-k I
kamlau Ix2	aliu III2	-jau Ii2
kamhel Ix3	alisæl III3	-taj Ii3
kamar Ixp	alia IIIp	-ja Iip

Anejom subjective particles once showed a four-way number distinction as well, but the system has been simplified in the past century. Tanna verbs mark four numbers-of-subject, but do so with preverbal prefixes which are morphologically distinct, and potentially physically separated, from the markers of person-of-subject.

In contrast, the focal, possessive and objective pronominal forms in Erromangan languages distinguish only two numbers, singular and plural. Thus the list of Sie focal pronouns given in (2) above represents *all* Sie focal pronouns, whereas the corresponding Anejom list can be filled out with four trial and four plural forms. However, Sie verbs overtly mark a singular/dual/plural distinction for first person (inclusive and exclusive) subject.

1.3 Case

Pronominal forms in all SV languages formally distinguish three cases - focal, possessive and subjective. Anejom, South-west Tanna and the Erromangan languages also formally distinguish objective pronouns from pronouns in each of the other cases. This four-way distinction is illustrated below:

(4)		SIE		S.W. TANNA		ANEJOM
	FOCAL	kam	Ixp	kəlau	Ii2	aek II
	OBJECTIVE	-ɣam	Ixp	alau	Ii2	-ɣ II
	POSSESSIVE	-mam	Ixp	-lau	Ii2	-m II
	SUBJECTIVE	kak-	Ixp	k-	Iins	nei II

The case roles are illustrated with Sie data below. *Focal* pronouns stand as emphatic subjects (cf. kam in (5)) and as NP arguments in some verbless constructions; *subjective* pronouns cross-reference the person (and, in some cases, the number) of the subject in verbs (cf. kak- in (5)):

(5)	kam kak-l-oyh-i	versus	kak-l-oyh-i
	Ixp Ixp-p-see-III		Ixp-p-see-III
	<i>We saw it/him</i>		<i>We saw it/him</i>

Objective pronouns cross-reference the person and number of the object of the verb (cf. -ɣam in (6)):

(6)	ɣ-okil-ɣam
	III-know-Ixp
	<i>He knew us</i>

Possessive pronouns are suffixed to nouns or to 'possessive markers' (cf. -mam in (7)):

(7)	noru-mam	nimo	hor-mam
	<i>hand-Ixp</i>	<i>house</i>	POSS-Ixp
	<i>Our hands</i>	<i>Our house</i>	

Of the Tanna languages, only South-west Tanna makes a formal distinction between focal and objective pronouns. The other languages use the focal pronoun to mark a pronominal object:

(8)		LENAKEL		KWAMERA
	FOCAL	} ilk II	}	iou I
	OBJECTIVE			
	POSSESSIVE	-m II		-k I
	SUBJECTIVE	n- II		ia- I

This lack of distinction between focal and objective forms is illustrated in the following Lenakel example:

(9)	iik n-əm-hoapwnín	kuri	taha-k!	i-ak-amwakín	iik!
	II	II-PST-kill:TR	dog	POSS-I.	I-CONC-hate:TR II
	<i>You</i>	<i>killed my dog!</i>	<i>I</i>	<i>hate you!</i>	

1.4 Internal structure

As will be clear from the examples given in (2) and (3) above, non-singular pronominal bases are not simply compound forms of the singular pronominal bases: each language formally distinguishes singular and non-singular pronominal forms of the same person and case.

However, in Anejom and the Tanna languages, which *do* distinguish more than one non-singular number, it is equally clear that non-singular pronominal forms are composed of a pronominal base and a number suffix. This can be clearly illustrated by an examination of the full set of Lenakel non-singular focal pronouns:

(10)	Ii	Ix	II	III
2	kalau	kamlau	kamilau	ilau
3	kathel	kamhel	kamhiel	ilhel
p	katar	kamar	kamiar	ilar

Given two phonological rules which are needed elsewhere in the language, it will be seen that the forms in (10) are compounds built according to the following principles:²

(11)	PRONOMINAL BASE	+	NUMBER MARKER
	kat-	Iins	-lau 2
	kam-	Ixns	-hel 3
	kami-	IIns	-ar p
	il-	IIIns	

In these languages also, neither pronominal bases nor number markers may stand alone: *both* are bound forms.

Given these facts, I will proceed by examining in section 2 the pronominal bases for the various persons and cases in the languages of the SV group, and in section 3 the markers of number; in each case, I will reconstruct PSV forms, and comment on their development from POC and/or Proto-Eastern Oceanic (PEO) etyma. I will make rather more general comments about the PSV system and its development in section 4.

2. PRONOMINAL BASES

This section will examine, in turn, the base forms of the focal, objective, possessive and subjective pronouns in the SV languages,³ on the basis of which the PSV base forms will be reconstructed, and will also examine the development of POC/PEO forms in PSV.

2.1 Focal pronouns

The full set of focal pronouns in the SV languages is given in Table 2. The Erromangan forms are monomorphemic, while the Tanna and Anejom forms consist of a pronominal base plus a number marker.

2.1.1 Reconstruction

The Anejom forms, however, show an additional complexity. As will be seen from an examination of Table 2, all Anejom focal pronouns begin with *a*. Now an NP subject in Anejom is marked by a proposed subject-marking particle *a*:

Table 2: Focal pronouns

	SIE	URA	UTH	NTN	WSN	LEN	SWT	KWM	ANJ
I	yau	yau	yo	iio	iəu	io	iou	iou	añak
II	kik	ga	ko	ik	ik	iik	iik	ik	æk
III	iyi	iyi	iyi	in	in	in	in	in	aen
Ii2				kilhao	kilau	kalau	kəlau	krau	akajau
Ii3				kitahal	kitahal	kathel	kətasəl	kətahar	akataj
Iip	koh	gis	gis	kitat	kitah	katar	kətaua	kətaha	akaja
Ix2				itlao	itəlau	kamlau	kəmlu	kəmrau	ajamrau
Ix3				itmahal	itəmahal	kamhel	kəmasəl	kəmrahar	ajamtaj
Ixp	kam	gim	kum	itmat	itəmah	kamar	kəmaua	kəmaha	ajama
II2				itəmlao	itəmwəlau	kamilau	kəmilu	kəmirau	ajourau
II3				itəmahal	itəmwahal	kamhiel	kəmisəl	kəmirahar	ajoutaj
IIP	kimi	gimi	kimi	itəmat	itəmwah	kamiar	kəmia	kəmiaha	ajowa
III2				ilao	ilau	ilau	iliu	irau	aarau
III3				ilahal	ilahal	ilhel	ilīsəl	irahar	attaj
IIIp	iror	leil	yoril	ilat	ilah	ilar	ilia	iraha	aara

- (12) et γiñ wametey a pikaθ
 III:AOR eat:TR sweet:potato SM pig
 The pig ate/is eating sweet potato
- et atŋii pikaθ a natamwāñ
 III:AOR kill:TR pig SM man
 The man killed the pig

Focal pronoun subjects, however, are not preceded by the particle a:

- (13) et γiñ wametey aen
 III:AOR eat:TR sweet:potato III
 He/she/it ate/is eating sweet potato
- *et γiñ wametey a aen
 III:AOR eat:TR sweet:potato SM III

Since focal pronouns most typically occur as subjects, it seems reasonable to suggest that the initial a on all Anejom focal pronouns is the fused subject-marker.

Given this analysis, plus a number of phonological rules for each language like those referred to above in 1.4, we can list the following pronominal base forms for each language:

- (14)
- | | SIE | URA | UTH | NTN | WSN | LEN | SWT | KWM | ANJ |
|--------|------|------|-------|-------|--------|-------|-------|-------|-------|
| I | yau | yau | yo | iio | iəu | io | iou | iou | -ñak |
| II | kik | ga | ko | ik | ik | iik | iik | ik | -ek |
| III | iyi | iyi | iyi | in | in | in | in | in | -en |
| Iins | koh | gis | gis | kit- | kit- | kat- | kət- | kət- | -kaj- |
| Ixns | kam | gim | kum | itm- | itəm- | kam- | kəm- | kəm- | -jam- |
| IIIns | kimi | gimi | kimi | itəm- | itəmw- | kami- | kəmi- | kəmi- | -jou- |
| IIIIns | iror | leil | yoril | il- | il- | il- | ili- | ir- | -ar- |

PSV reconstructions for each person will be discussed below. Although I have not finalised my reconstruction of PSV proto-phonology, I will comment on phonological matters where appropriate.

FIRST SINGULAR. The Anejom form is not cognate; it will be discussed below in section 2.3. The other forms point to a PSV reconstruction *iau, which presumably derives from POC *i-au (personal:article-I).⁴

SECOND SINGULAR. The proposed reconstruction is *ig(ao), which presumably derives from POC *i-ko(e) (personal:article-II). The Tanna languages and Anejom develop regularly,⁵ showing normal PSV final vowel loss (Lynch 1978a:739). The Erromangan languages, however, show some irregularities: Sie shows accretion of an initial k, while Ura and Utaha show irregular loss of initial *i and unexpected retention of the final vowel.

THIRD SINGULAR. Tanna and Anejom suggest a reconstruction *in (or, more correctly, *in(V), since if a final vowel were present we would expect it to be lost; this final vowel could not have been *i or *e since *n becomes ANJ ñ in this environment).⁶ In Erromangan languages, y reflects both POC *i and *y and also, in one or two cases, POC *ñ: POC *ñamuk *mosquito* > SIE u/yomuɣ, URA yomoy, for example. It is tempting to hypothesise that the ancestral third person pronoun was derived from a form such as PEO *inia, which was inherited as *iña. Since the commonest reflex of *ñ appears to be n in Tanna and Anejom, the forms there are regular; on the other hand, had the development *ina > *iya taken place in Erromango, subsequent raising of *a > i may well have occurred (cf. POC *kamiu IIns > SIE kimi, *natu *child* > SIE nit-, *mate *die* > URA mis, *kayu *tree* > URA n/ii as examples of the raising of *a to i in various environments). It is not yet clear, however, that reconstruction of a palatal nasal is warranted for PSV. For the moment then, we reconstruct PSV *in(V) on the basis of Tanna and Anejom only, treating the Erromangan forms as exceptions of a kind.

FIRST INCLUSIVE NON-SINGULAR. The forms given in (14) suggest a PSV reconstruction *gadi, which is presumably a metathesised form of POC *kinta. Both Anejom and the Erromangan languages show palatalisation of *d, typical of dento-alveolar stops before a high vowel.⁷ Further, although the i in Ura and Utaha gis can be explained by the rule of *a-raising discussed in the previous paragraph, the vowels in Sie koh, Lenakel kat-, Anejom -kaj-, etc. could not be explained as deriving from *i: hence, the metathesis seems established.

The Tanna forms require some discussion. First, the reflex of *d is a stop, which suggests that while POC *t underwent palatalisation before *i, POC *nt did not; unfortunately, there are no other examples of POC *nti sequences reflected in these languages which would allow us to test this hypothesis. Second, while Lenakel, South-west Tanna and Kwamera clearly reflect the metathesis, North Tanna and Whitesands both have kit-. The rule of *a-raising is less common in these languages than in those of Erromango, although there are a few examples; e.g. *kali *dig* > NTN il (but cf. WSN el); *masakit *sick* > NTN *misaŋ. This suggests, then, that all forms derive from *gadi.

FIRST EXCLUSIVE NON-SINGULAR. Two quite distinct reconstructions need to be made here. SIE kam, URA gim, UTH kum, LEN kām-, and SWT, KWM kām- all derive from a form *gami (which is a reflex of POC *kami). The final *j, before being lost, conditions the raising of *a in Ura. Only the vowel of the Utaha form remains unexplained, and Utaha data are in any case not reliable.

On the other hand, NTN itm-, WSN itəm-, and ANJ -jam- are clearly not derived from *gami. The Tanna languages suggest a reconstruction *idam(V), which will also explain the Anejom form if *d palatalised adjacent to *i (and not just before *i, as discussed above). I have at this stage no additional evidence which would confirm or deny this.

SECOND PERSON NON-SINGULAR. Again, two reconstructions are required, and the implications of two reconstructions for first person exclusive and second person non-singular will be discussed below. *gamiV (< POC *kamiu) has reflexes in the Erromangan languages (where *a-raising has taken place), Lenakel, South-west Tanna and Kwamera. The final unspecified vowel must have been present, since otherwise the final vowel deletion rule would have deleted the *i. This final vowel was probably *u (given the nature of the POC form), but the data do not allow us to establish this; however, since in 2.2.1 I reconstruct *-yamiu for 'Iins objective', I think it is safe to reconstruct final *u.

The second reconstruction, on the basis of North Tanna, Whitesands and Anejom data, is *idamu(V). The problem regarding palatalisation of *d and j in Anejom, which was discussed above, occurs again. Anejom also shows unexplained loss of *m. It is necessary to reconstruct *u to explain both Anejom u and the velarised nasal in Whitesands. The final vowel may not have been present, since although Anejom generally deletes final vowels, a final vowel cluster is not affected (e.g. POC *dua two > ANJ e/rou).

THIRD PERSON NON-SINGULAR. It seems clear that the basic form of this pronoun is *iL.⁸ The vowel of the Anejom form is irregular, the expected form being -er-. However, just as aek II and aen III are frequently heard as [a:k] and [a:n] - i.e. the e assimilates to the a of the subject marker - so underlying /a-er-/ may well have undergone the same assimilation, yielding a-ar-. The Erromangan forms show what appears to be some kind of reduplication, with possible metathesis as well in Ura; in each language the two vowels of the root differ, suggesting that the reduplication was only partial.

The basic form of the PSV third non-singular focal pronoun was reconstructed as *iL. That form must have had a supporting vowel, since *L is lost finally in Anejom. The South-west Tanna form ili- suggests that this vowel should be reconstructed as *i, and since no other forms contradict this - apart from the Erromangan forms, which are contradictory in other ways - I reconstruct PSV *iLi.

SUMMARY. The PSV focal pronominal bases which have been reconstructed here are listed below:

(15)		*gadi	Iins	
	*iau	I	*gami, *idam(V)	Ixns
	*ig(ao)	II	*gamiu, *idamu(V)	IIns
	*in(V)	III	*iLi	IIIns

2.1.2 Development

Of the forms listed in (15), only those reconstructed for I and III appear not to require further comment, deriving respectively from POC *i-au and a PEO-like form *inia.

The remaining non-third person forms contain the proto-phoneme *g, which requires some comment. The 'normal' reflex of POC, PEO *k is PSV *k.⁹ PSV *g

appears to be restricted in its distribution to pronouns (including those listed above), a numeral prefix *gV-, and possibly one or two other items. In this it recalls the restricted distribution of Raga k (Walsh 1982); although a number of details are different, a similar statement can be made - the reflex of *k in pronouns (a) is rare and (b) is phonemically different from the reflex of *k in the vast majority of lexical items.

With this comment, which bears considerable further investigation but which is beyond the scope of this paper, we can dispose of PSV *ig(ao) II (< POC *ko(e), PEO *i-koe), *gami Ixns (< POC, PEO *kami), and *gamiu IIns (< POC, PEO *kamiu).

Of the remaining forms, both *gadi Iins and *ili IIns show irregular developments: the former is a clear case of metathesis from POC *kinta, while the latter shows final *i for expected *a (the forms *sida, *kida and *ida have been reconstructed by different scholars for POC and/or PEO; but see the discussion under 2.3.2). While the occurrence of *g in four pronominal forms suggests clear links with (some) languages to the north, these two innovations appear to be exclusive to the SV subgroup.

Finally, the forms *idam(V) Ixns and *idamu(V) IIns, reconstructed on the basis of evidence from Anejom and two of the Tanna languages, do not, as far as I am aware, have cognates elsewhere in Oceanic. I shall return to these forms in section 4.

2.2 Objective pronouns

The full set of objective pronouns in those SV languages which formally distinguish these from the focal pronouns is given in Table 3.

Table 3: Objective pronouns

	SIE	URA	UTH	SWT	ANJ
I	-yau	-yau	-yo	[same	-ñak
II	-y	-ka	-ko	as	-yiγ, -γ
III	-i, -∅	-i, -∅	-i	focal]	-yin, -n
Ii2				alau	-γajau
Ii3				atasəl	-γataj
Iip	-yoh	-gis	-kis	ataua	-γaja
Ix2				amlu	-γamrau
Ix3				amasəl	-γamtaj
Ixp	-yam	-gim	-kum	amaua	-γama
II2				amilu	-γourau
II3				amisəl	-γoutaj
IIP	-yum	-(n)imi	-kimi	amia	-γowa
III2				aliu	-rau
III3				alisəl	-ettaj
IIIP	-or	-(i)l	-kor	alia	-ra

2.2.1 Reconstruction

It will be immediately apparent from Table 3 that some languages do not formally distinguish focal and objective pronouns in some persons, using instead the focal forms: viz. Sie in I, Ura in I, Iins and Ixns; Utaha in I, II, Ixns and IIns; and South-west Tanna in I, II and III. These forms are not classed as 'true' objective pronouns.

Table 3 also shows some variant forms. Of the two Sie and Ura forms for III, -i is the base form, -∅ being used after verb roots which end in i. Of the two Anejom forms for II and III, -yiɣ and -yin occur after consonant-final roots, while -ɣ and -n occur after vowel-final roots: I will take the shorter forms as being basic, since the longer forms may well derive from transitive suffix -i + pronoun.

We can thus list the 'true' objective pronominal bases as follows:

(16)	SIE	URA	UTH	SWT	ANJ
I	-	-	-	-	-ñak
II	-ɣ	-ka	-	-	-ɣ
III	-i	-i	-i	-	-n
Iins	-ɣoh	-	-kis	at-	-ɣaj-
Ixns	-ɣam	-	-	am-	-ɣam-
IIns	-ɣum	-(n)imi	-	ami-	-you-
IIIIns	-or	-(i)l	-kor	ali-	-r-

The singular forms can be easily disposed of. In the first singular, the only 'true' objective pronoun is Anejom -ñak, but I treat this as derived from the possessive, for reasons discussed below in 2.3; thus no reconstruction can be made. In the second person, Sie and Anejom suggest *-ɣ(V), the Ura form -ka quite possible being a weakened form of the focal pronoun ga. In the third person, the Erromangan forms are homophonous with the transitive suffix, and may not be 'true' pronouns, while Anejom -n may be a reduced form of focal aen; in any case, no reconstruction can be made.

In the non-singular, the first person forms show a clear relationship with the focal forms, the major difference being the weakening of the initial consonant from *g to *ɣ; I thus reconstruct *-ɣadi Iins and *-ɣami Ixns. The second person forms vary rather more significantly: note particularly the vowel in Sie -ɣum and the second vowel in Anejom -you, which suggest the presence of *u in the ancestral form. I thus tentatively reconstruct *-ɣamiu II (cf. focal *gamiu) to explain both Sie and Anejom u and Ura and South-west Tanna i. What is of interest here is that, although the Anejom first exclusive and second person non-singular focal forms derive from the non-Oceanic roots *idam(V) and *idamu(V), the corresponding objective (and, as we will see, possessive) forms appear to be regular developments. Finally, the third person non-singular form derives from a reconstruction *-Vli with an unspecified initial vowel; Utaha k is not explained.

The PSV objective pronominal bases reconstructed above are listed again below:

(17)		*-ɣadi	Iins
?	I	*-ɣami	Ixns
*-ɣ(V)	II	*-ɣamiu	IIns
?	III	*-Vli	IIIIns

2.2.2 Development

The innovations of metathesis in the form for Iins and irregular final vowel developments in the form for IIins, noted above with respect to the focal forms, recur in the objective forms, and require no further comment. There is little else to discuss here, apart from the fact that those pronouns which showed *g in the focal form show a corresponding *γ in the objective form - again, not the regular reflex of POC/PEO *k.

2.3 Possessive pronouns

The full set of possessive pronouns in all SV languages - including Sorung, for which there are data available - is given in Table 4.

2.3.1 Reconstruction

It should be noted that the Erromangan forms listed in Table 4 are those which are suffixed to nouns in direct (or inalienable) possessive constructions or to the possessive morphemes Sie hor-, Sorung sor-; Sie suffixes focal pronouns to the possessive morpheme en- (Lynch and Capell 1983:43-44), while Ura shows a considerable breakdown of what appears to have been an earlier, more typical Oceanic, possessive system (Lynch 1983b:154) - only the vestiges of the earlier Ura system are given in Table 4.

Table 4: Possessive pronouns

	SIE	URA	UTH	SOR	NTN	WSN	LEN	SWT	KWM	ANJ
I	-ŋ	-k	-ŋ	-ŋ	-k	-k	-k	-k	-k	-k, -ñak
II	-m	-m	-ko	-m	-m	-m	-m	-m	-m	-m
III	-n	-n	-n	-n	-n	-n	-n	-n, -ni	-n, -ni	-n
Ii2					-lhao	-lau	-lau	-lau	-rau	-jau
Ii3					-tahal	-tahal	-thel	-tasəl	-tahar	-taj
Iip	-nt	?	?	-t	-tat	-tah	-tar	-taua	-taha	-ja
Ix2					-tlao	-təlau	-mlau	-mlu	-mrau	-mrau
Ix3					-tmahal	-təmahal	-mhel	-masəl	-mrahar	-mtaj
Ixp	-mam	-kim	-kum	-mam	-tmat	-təmah	-mar	-maua	-maha	-ma
II2					-təmlao	-təmwəlau	-milau	-milu	-mirau	-mirau
II3					-təmahal	-təmwəhal	-mhiel	-misəl	-mirahar	-mitaj
IIP	-mi	?	?	-mi	-təmat	-təmwəh	-miar	-mia	-maha	-mia
III2					-lao	-lau	-nilau	-liu	-nrau	-rau
III3					-lahal	-lahal	-nilhel	-lisəl	-nrahar	-ttaj
IIIP	-d	?	-ira	-da	-lat	-lah	-nilar	-lia	-nraha	-ra

Of the data given in Table 4, the following appear not to be 'true' possessive suffixes, but rather (modifications of) focal or objective pronouns: Ura -kim Ixns, Ura -ko II and -kum Ixns. The base forms of the remaining possessive forms for which data are available are:

(18)	SIE	URA	UTH	SOR	NTN	WSN	LEN	SWT	KWM	ANJ
I	-ŋ	-k	-ŋ	-k	-k	-k	-k	-k	-k	-k, -ñak
II	-m	-m		-m	-m	-m	-m	-m	-m	-m
III	-n	-n	-n	-n	-n	-n	-n	-n, -ni	-n, -ni	-n
Iins	-nt			-t	-t-	-t-	-t-	-t-	-t-	-j-
Ixns	-mam			-mam	-tm-	-təm-	-m-	-m-	-m-	-m-
IIns	-mi			-mi	-təm-	-təmw-	-mi-	-mi-	-mi-	-mi-
IIIns	-d		-ira	-da	-l-	-l-	-nil-	-li-	-nr-	-r-

Of the variations noted in (18), South-west Tanna uses -ni III after (most?) kinship terms and -n elsewhere (Lynch 1982a:28). Kwamera appears to use -ni III after most (all?) kinship terms and also after some possessive morphemes. Anejom uses -ñak I after the 'neutral-active' possessive morpheme u- and -k elsewhere (Lynch 1982b:123ff).

A number of reconstructions of PSV possessive suffixes are problematical. Those which are not include the following:

(i) There appears to be clear agreement on the reconstruction of PSV *-m(V) II.

(ii) We can reconstruct PSV *-n(V) III; the contrast between South-west Tanna and Kwamera -n and -ni may indicate a PSV distinction between two kinds of direct possession (cf. Lynch 1982d), but until we have more detailed data from both these languages and other Oceanic subgroups, we cannot be sure about this.

(iii) We can reconstruct PSV *-mam(V) Ixns on the basis of the Erromangan forms and the POC form *-mami. Anejom and the Tanna languages appear to have undergone reduction, to a form deriving from *-m(V). (The North Tanna and Whitesands forms for both Ixns and IIns derive from the 'non-Oceanic' focal pronouns *idam(V) and *idamu(V) and will not be discussed here.)

(iv) The reconstruction PSV *-miV IIns is clear, the final *V being required to protect the preceding *i from the final vowel loss rule.

(v) In reconstructing the PSV form *-Lia IIIns, I treat as an accretion the apparent prefixation of the third person singular suffix in the Lenakel and Kwamera forms. The vowels of this suffix require brief comment: *i is overtly reflected in South-west Tanna, and its presence accounts for the liquid (rather than high front vowel) reflex of *L in the three northern Tanna languages; while *a is reflected in Utaha and Sorung (although we might have expected this to be deleted by the final vowel loss rule).

We are thus left with the first person singular and first person inclusive non-singular forms to discuss. I will examine the singular form first.

As mentioned above, Anejom shows two allomorphs of the form for I, -k and -ñak; the latter also occurs as an objective suffix and a focal pronominal base, and yet it does not appear to derive from any reconstructed Oceanic first person singular pronominal form. Where does Anejom -ñak come from? The only suggestion that I can make at this stage is that it is a borrowing from the Polynesian Outlier language Futuna-Aniwa. According to Capell (1960:105ff), Futuna-Aniwa has a predicative possessive first person singular form niaku *mine* (probably morphologically ni 'predicative' + a 'dominant' + ku I). There has been considerable Futunese migration to Aneityum over the past two centuries, and Anejom is a language which can be demonstrated to have undergone considerable change in the last century at least (cf. 2.4 below and also Lynch 1982c). The development of Futuna-Aniwa niaku > Anejom -ñak is clearly phonologically plausible;

the borrowing of the form in the first place, and its subsequent generalisation to focal and objective cases is clearly less plausible, but not outside the bounds of possibility.

We are left with the problem of nasal reflexes in Sie, Utaha and Sorung corresponding with stop reflexes in Tanna and Anejom. The Tanna and Anejom forms would be consistent with either of the PSV reconstructions *-k(V) or *-g(V). Ura -k at first glance suggests the former reconstruction, since PSV *g > Ura g while PSV *k > Ura k; however, since we have little evidence for *g in any case, and virtually none in final position, the Ura form might well be a reflex of *-g(V).¹⁰ The -ŋ of the remaining Erromangan languages may derive from the breaking and weakening of a prenasalised stop (i.e. *g = [ŋg] > ŋg > ŋ/___ #), or it may derive from the weakening of a nasal + stop cluster (supporting a reconstruction *-ŋk(V)). On balance, I believe the evidence favours the reconstruction of *-g(V).¹¹

The first person inclusive non-singular possessive suffix also poses problems. Recall that a focal form *gadi and an objective form *yadi were reconstructed above, in part to explain the palatal reflexes of the second consonant in Anejom and the Erromangan languages. The coronal consonant behaves in the same way in the possessive in Tanna and Anejom as it does in the focal and objective, and these forms suggest a reconstruction *-di. The Erromangan languages, however, do not show the expected palatalisation in this form, and suggest a reconstruction *-d(V) (or possibly *-nt(V)), where *V is not *i. I will attempt to explain these two forms below; however, since the former is found in two subgroups, while the latter probably continues POC *-nta, I reconstruct both at this stage.

The PSV possessive pronominal bases which were reconstructed above are listed below:

(19)			*-di, *-dV	Iins
	*-g(V)	I	*-mam(V)	Ixns
	*-m(V)	II	*-miV	IIIns
	*-n(V)	III	*-Lia	IIIns

2.3.2 Development

The possessive suffixes reconstructed for POC and PEO are given in (20):

(20)	PEO	PEO		POC	PEO	
				*-nta	*-(n)ta	Iins
	*-ŋku	*-(ŋ)ku	I	*-mami	*-mami	Ixns
	*-mu	*-mu	II	*-miu	*-m(i)u	IIIns
	*-ña	*-ña	III	*-da	*-nda	IIIns

A comparison of (20) with (19) indicates that there is little need to comment on the development of the PSV possessive suffixes, except for the forms for Iins and IIIns.

The form *-Lia IIIns again shows unexpected *i. There are other occurrences of an intrusive i in this form in other branches of Oceanic - the Central Papuan languages, for example, where Ross (1983) reconstructs *yidi(a) 'IIIp focal', *-di(a) 'IIIp possessive' - which suggest that the PSV forms are not all that unusual.

The two forms for Iins require further comment. The form *-d(V) was probably *-d(a), and historically reflects the POC and PEO forms. The form *-di, on the other hand, is clearly phonologically related to the metathesised focal and objective forms *gadi and *-yadi. The phonological connection between POC non-singular possessive pronouns on the one hand, and focal and objective pronouns on the other, is fairly transparent: drop the first syllable of the focal form and you are left with the possessive form. Following this 'rule', the metathesis *kinta > *gadi could have led to a substitution of *-di for expected *-da. The Erromangan languages, on the other hand, appear to have 'ignored' the phonological relationship and retained *-nta as *-da. The form *-di thus represents an innovation based on analogy: whether it is sufficient to unite Tanna and Anejom as a single subgroup as against Erromangan I tend to doubt, since it is the kind of innovation which could have occurred independently.

2.4 Subjective pronouns

2.4.1 Reconstruction

The situation with regard to the subjective pronouns is quite different from that relating to pronouns in the other three cases. It will be useful to examine the forms involved subgroup by subgroup, and then list the base forms in a complete table.

ERROMANGO. Data on subjective pronouns in Ura and Utaha are scarce, and in Sorung non-existent. I have interpreted what Ura and Utaha data are available in the light of what is known about Sie subjective pronouns. Two general points can be made here. First, markers of person may be separated from markers of non-singular number by tense/aspect markers; e.g.:

- (21) ko-m-l-oyh-i
 Iins-MP-p-see-III
We saw it

Second, Sie shows two phonologically distinct sets of person-of-subject markers to verbs: one, Set A, is used with verbs in the general past, present and future tenses; while the other, Set B, is used with the mid past, far past, past conditional and non-past conditional tenses. The Sie forms for both sets are given below. Since the same kind of system appears to operate in Ura, I have given available forms according to Set membership in that language as well. The Utaha data are insufficient to allow us to draw any conclusions about Set A/B distinctions.

(22)	SIE		URA		UTAHA
	Set A	Set B	Set A	Set B	
I	yay-	yau-	yay-	yau-	i-
II	k-	ki-	k-	k-	k-
III	γ-	γ-	γ-	γ-	∅-(?)
Iins	kok-	ko-	ku-		
Ixns	kak-	ka-	kak-		ku-
IIins	k-	ki-	ki-		ki-
IIIins	γ-	d-	il-	il-	el-

As a comparison with Table 2 will show, the Set B forms appear to be clearly derivable from the focal pronouns - the classical 'Melanesian short pronouns'. I thus take the Set A forms as being basic, and will not refer to the Set B forms again.

TANNA. As in Erromango, person and number of the subject are marked by distinct prefixes which may be separated from each other by other prefixes. The following South-west Tanna verb illustrates this:

- (23) i-akn-am-ha-vən
Ixns-INCH-CONT-p-go
We have begun to go

However, unlike the Erromangan situation, the Tanna languages use the same set of pronominal forms with all tenses and aspects. These are:

(24)	NTN	WSN	LEN	SWT	KWM
I	i-	i-	i-	i-	ia-
II	n-	n-	n-	n-	i-
III	t-	t-	r-	l-	r-
Iins	k-	k-	k-	k-	k- (dual), s- (elsewhere)
Ixns	i-	i-	i-	i-	ia-
IIns	n-	n-	n-	n-	i-
IIIIns	k-	k-	k-	k-	k- (dual), Ø- (elsewhere)

ANEJOM. The situation here is different again, and more complex. Inglis (1882) and Capell (n.d.) list a set of portmanteau preverbal particles, partly analysable, marking person and number of subject and tense/aspect. It is clear, however, that there has been massive simplification of the system within the last century: the forms which I recorded show neutralisation of the dual/trial/plural distinction, as well as levelling of some person distinctions. The full set of forms for three tense/aspects in both 'Old' and 'Modern' Anejom is given in Table 5.¹²

In this historical study, I take the 'Old' Anejom forms as the basis for comparison with other SV languages. Although there appear to be a number of conflationary phonological rules involved, some of them somewhat idiosyncratic, this is only to be expected in particles of such high frequency. The aorist forms are the simplest, the non-singular consisting of a pronominal base + number-markers of the form -au '2', -taj/-tij '3', and -a 'p'. The past forms are derivable from the aorist form + is, while at least some of the subjunctive forms seem to be derivable from the aorist form + i. This suggests that the following forms can be isolated as underlying markers of person-of-subject in 'Old' Anejom:

(25)		int-/t-	Iins
ek/k-	I	eyr-	Ixns
na/n-	II	ek-	IIns
et/y-/Ø-	III	er-	IIns

Of these, k-, n-, and t- are clearly abbreviated forms of ek, na and int-, and will be ignored for the remainder of this discussion.

Table 5: 'Old' and 'Modern' Anejom subjective pronouns

	AORIST		PAST		SUBJUNCTIVE			
	'Old'	'Modern'	'Old'	'Modern'	'Old'	'Modern'		
I	ek	ek,k-	kis	kis, is	inki,ki	ki		
II	na	nei,na-	as	is	an	ni		
III	et	et,t-	is	is	inyi,yi	iñiyi,yi		
Ii2	intau	all persons and numbers show	intis	kis, is	tu	ri		
Ii3	intaj		intijis		tiji			
Iip	inta		imjis		ti			
Ix2	eyrau		eyrus	is	eyru			
Ix3	ektaj,ektij		ektijis		tiji			
Ixp	eyra		eyris		eyri			
II2	ekau		ekra,	akis	ekris, is		eru	ri, ra
II3	ahtaj		era,	ahtijis			tiji	
Iip	eka		rai-	akis			aki	
III2	erau			erus	eru			
III3	ehtaj		ehtijis	tiji				
IIIp	era		eris	eri				

We can now summarise this information by giving the base forms of the subjective pronouns in all the SV languages. This summary is provided in Table 6. Ura and Utaha forms in parentheses may well be Set B forms.

Table 6: Subjective pronouns

	SIE	URA	UTH	NTN	WSN	LEN	SWT	KWM	ANJ
I	yay-	yay-	i-	i-	i-	i-	i-	ia-	ek
II	k-	k-	k-	n-	n-	n-	n-	i-	na
III	ɣ-	ɣ-	∅-	t-	t-	r-	l-	r-	et/ɣ-/∅-
Iins	kok-	(ku-)	?	k-	k-	k-	k-	k-/s-	int-
Ixns	kak-	kak-	(ku-)	i-	i-	i-	i-	ia-	eyr-
IIns	k-	(ki-)	(ki-)	n-	n-	n-	n-	i-	ek-
IIIins	ɣ-	il-	(el-)	k-	k-	k-	k-	k-/∅-	er-

The second person forms are perhaps the easiest to deal with. Sie does not distinguish II from IIns, marking both with k. The Tanna languages likewise do not distinguish II from IIns; a Proto-Tanna reconstruction *n(V) can be posited, Kwamera showing an innovation. Anejom, however, does distinguish II from IIns, the singular form probably cognate with the Tanna form and the plural form with Sie. I hypothesise that PSV marked II with *na- and IIns with *k(V)-, that Sie (and the other Erromangan languages?) generalised the non-singular form to the singular, and that Proto-Tanna generalised the singular form to the non-singular.

Of the remaining forms, I will leave discussion of the first singular reconstruction until last, as it is perhaps the most complicated. On the basis of the available evidence, no reconstruction can be made for Ixns, each subgroup showing an historically different form. The remaining reconstructions are reflected in two of the three subgroups only. First, *t(V)- III is reconstructible

on the basis of Tanna and Anejom evidence only;¹³ the Erromangan form γ - may be another case of generalisation from non-singular to singular. Second, I propose the reconstruction $*\gamma(V)$ - for IIIIns on the basis of Sie and Tanna evidence:¹⁴ PSV $*\gamma$, though generally lost in South-west Tanna and Kwamera, is occasionally reflected (as k), and its loss in this morphological environment would make IIns tenseless verbs indistinguishable from imperatives; note in any case the grammatically conditioned variation between k and \emptyset in Kwamera. Finally, the forms for IIns in Sie and Tanna suggest a reconstruction $*k(V)$ -, or perhaps $*k[ok]$ - or $*k[ok](V)$ -.

We return now to the first person singular. The Sie, Ura and Anejom forms suggest a reconstruction something like $*iaK(V)$ - where $*K$ represents a velar of some kind which I will not specify at this stage. Kwamera shows loss of the velar (which would be expected if the velar were $*\gamma$), but the other Tanna languages show even further contraction, retaining only the high vowel. I believe the full reconstructed form can be supported, and that there has been morphological reanalysis in Tanna. I will, for the sake of convenience, use only Sie, Lenakel and Kwamera examples to illustrate this point.

Sie has a present tense marker am -, which follows the subject marker. The behaviour of the relevant prefixes is illustrated below; the apparently unmarked form (26a) is in fact marked as irrealis by nasalisation of the root (cf. Lynch and Capell 1983:23-25).

- | | |
|---------------------------------------|--|
| (26) a. $yay\text{-}a\eta h\text{-}i$ | b. $yay\text{-}am\text{-}a\eta h\text{-}i$ |
| I-IRR: <i>see-it</i> | I-PRES-IRR: <i>see-it</i> |
| <i>I will see it</i> | <i>I see it</i> |

Both Lenakel and Kwamera have, save for the change from γ to k, verb-initial sequences which are virtually identical to (26b):

- | | |
|---|---------------------------------------|
| (27) LEN: $i\text{-}ak\text{-}am\text{-}ol$ | KWM: $ia\text{-}k\text{-}am\text{-}o$ |
| I-CONC-CONT- <i>do</i> | I-CONC-CONT- <i>do</i> |
| <i>I am doing it</i> | <i>I am doing it</i> |

Note, however, that although the *phonology* of the first two syllables of (26b) and (27) is virtually identical (Sie [yayam], Lenakel and Kwamera [yagam]), the *morphological* structure is different in each case.

I propose that the original form of the prefix was $*iak(V)$ -. This form has been inherited into Anejom, and also - with the not unnatural change of $*k > \gamma$ in an environment which is intervocalic and usually unstressed - in Sie and Ura. I also suggest that a present continuous marker $*am$ can be reconstructed for PSV. The Erromangan languages retain the original PSV system; Anejom retains the pronoun, but not the tense marker. The Tanna languages, however, appear to have reanalysed the phonological sequence, with am being restricted to marking continuous aspect, while a part of the pronoun took on the function of marking present tense (called 'concurrent' due to various other features of this marker which need not concern us here). The sequence of developments was probably something like:

- | | | | | | |
|------|--------------------------|---|---------------------------------|--|---------------------------------|
| (28) | PSV | | KWM | | LEN |
| | $*iak\text{-}am\text{-}$ | > | $ia\text{-}k\text{-}am\text{-}$ | | $i\text{-}ak\text{-}am\text{-}$ |
| | I-PRES- | | I-PRES-CONT- | | I-PRES-CONT- |

The reconstruction hypotheses proposed above are summarised below:

(29)		*k(V)-	Iins
	*iak(V)-	I	?
	*n(V)-	II	*k(V)-
	*t(V)-	III	*γ(V)-
			Ixns
			IIIns
			IIIIns

If indeed the Iins form was *k(V)- and not *k[ok](V)- (as discussed above), then presumably the vowels of *k(V)- Iins and *k(V)- IIIns would have been different. Those vowels, however, are unrecoverable.

2.4.2 Development

It is immediately clear that the PSV subjective pronouns listed in (29) do not derive from the focal pronouns of either PSV or POC - i.e. they are not merely 'short pronouns' but a phonologically distinct set of subject-marking prefixes.

Ross (1981) has reconstructed a set of POC subject-marking prefixes. These are reproduced below in the form in which they are given in a more recent paper (Ross 1982:31):

(30)		*ta	Iins
	*ŋku	I	*mi,*ma
	*u,*ko	II	*m(i)u
	*i,*e	III	*(n)t(ie)
			Ixns
			IIIns
			IIIIns

A number of these forms recall POC focal forms or forms in other cases. What is more to the point, apart from a vague resemblance between PSV *iak(V)- and POC *ŋku I, none of the PSV forms appears to derive from the POC forms.

The PSV system, and the forms themselves, may thus constitute an innovation. However, considerably more research - particularly into the systems of Northern Vanuatu, the Loyalty Islands and New Caledonia - will be necessary before such a statement can be justified.

2.5 Summary

The reconstructed PSV pronominal bases for each case are listed together in (31) for convenience of reference:

(31)	FOCAL	OBJECTIVE	POSSESSIVE	SUBJECTIVE
	I	*iau	?	*-g(V)
	II	*ig(ao)	*-γ(V)	*iak(V)-
	III	*in(V)	?	*-m(V)
				*n(V)-
	Iins	*gadi	*-γadi	*-n(V)
	Ixns	*gami,*idam(V)	*-di,*-d(V)	*t(V)-
	IIIns	*gamiu,*idamu(V)	*-γami	*k(V)-
	IIIIns	*ili	*-mami(V)	?
			*-mi(V)	*k(V)-
			*-VLi	*γ(V)-
			*-Lia	

3. NUMBER-MARKING

As has been mentioned earlier, the SV languages mark non-singular numbers of the subject in various ways, while the languages of Tanna and Anejom also mark non-singular numbers in focal, objective and possessive cases by suffixes to the pronominal base. In this section, I investigate the form of these markers of number. Before doing so, however, it will be necessary to pay brief attention to the numerals in the SV languages and to other formal considerations.

3.1 Formal considerations

3.1.1 Numerals

The SV number-markers, as we will see, are phonologically *related* to the numerals - specifically, the dual, trial and plural markers bear a phonological relationship to the numerals *two*, *three* and *four*. The forms of these numerals are given below; forms in square brackets will not concern us here.

(32)	<i>two</i>	<i>three</i>	<i>four</i>
SIE	duru	dehel	dvat
URA	gelu	gehli	[lemelu]
UTH	kalu	kihili	[lemelu]
NTN	kəiu	kəsəl	kuvət
WSN	kəiu	kəsəl	kuvət
LEN	kiu	kəsil	kuvər
SWT	kəlalu	kəsisəl	kuas
KWM	kəru	kahar	kefa
ANJ	erou	esej	[mijman]

Of the numerals which are of concern to us, all show an identifiable numeral prefix: d(V)- in Sie, ge- in Ura, e- in Anejom, and k(V)- in the other languages. Note also that the South-west Tanna forms for *two* and *three* show partial reduplication.

The reconstructed PSV numerals are *-Lua *two*, *-s(ie)li *three*, and two (competing?) forms for *four*, *-vati and *-vat(V) (in which the final *V, if present, was not *i). These forms all clearly derive from POC reconstructed etyma - POC *dua *two*, *tolu *three*,¹⁵ and *pati or *pat *four*, the last two probably being doublets - and require little discussion here. Briefly, the final *a in *-Lua is justified by the way Anejom treats *ua sequences (cf. POC *luaq *vomit* > ANJ a/lou), while the final *i in *-s(ie)li is justified both by the nature of the reflex of *l in the northern Tanna languages and its (unexpected?) retention in Ura and Utaha. South-west Tanna and Kwamera derive their forms for *four* from *-vati, while the forms in the other languages just as clearly derive from the reconstruction *without* final *i.

3.1.2 Morphophonemic and other variation

There are considerable differences between forms marking number-of-subject and those marking number in focal, objective and possessive cases. First, Sie - and possibly the other Erromangan languages - has overt verbal prefixes marking dual and plural subject, but does not overtly mark pronominal number in any other

case. Second, Anejom and the Tanna languages mark the dual/trial/plural distinction in subjects by prefixed or preposed particles, but in other cases by suffixes. Finally, the phonological forms of subjective number-markers are not always identical to the phonological forms of non-subjective number markers: as one example only, Lenakel marks dual in non-subjective pronouns by the suffix -lau, but dual subject by one of the two allomorphs u- or ia-.

Table 7: Number-markers

	SIE	NTN	WSN	LEN	SWT	KWM	ANJ
2 Subj.	u-	u-,ia-	u-,ia-	u-,ia-	u-,la-	rou-, rau-,r-	-au
Non-Subj.	-	-lao	-lau	-lau	-lau, -lu,-u	-rau	-rau
3 Subj.	-	hl-	hal-	hal-, hai-	lh-	har-	-taj
Non-Subj.	-	-ahal	-ahal	-hel	-asəl, -səl	-rahar	-taj
p Subj.	u-,l-	ot-	ot-,oh-	ar-,ai-	ha-,s-	ha-	-a
Non-Subj.	-	-at	-ah	-ar	-aua,-a	-aha	-a

Nevertheless, despite these differences, I believe that a single set of number-markers can be reconstructed. Table 7 lists subjective and non-subjective number-markers in all SV languages for which sufficient data are available. Variants predicted by regular, independently motivated phonological rules are excluded; all other variant forms are included in Table 7, and these I will now briefly discuss before moving on to reconstruct each marker.

There are two cases where the allomorphic variation is grammatically conditioned. First, Sie makes a singular/dual/plural distinction in the first person, but only a singular/plural distinction in the other persons; further, the verbal prefix u- marks a first person subject as *dual* but a non-first person subject as *plural*, while a first person *plural* subject is marked with l-. Second, the variation in the Kwamera dual subject marker appears to be conditioned by the tense/aspect marker that follows: basically, r- occurs before perfective, sequential or conditional markers, while rou- (becoming rau- before some verbs) occurs elsewhere.

All other variation is phonologically conditioned. With regard to variation in the form of the dual subject-markers, the following general pattern obtains:

(33)		NTN	WSN	LEN	SWT
	before a, e, ə	u-	u-	u-	u-
	before o	ia-	ia-	ia-	u-
	elsewhere	ia-	ia-	ia-	la-

while in the plural, we can make the following generalisations:

(34)		WSN	LEN	SWT
	before a	ot-	ai-	s-
	before mid V	ot-	ar-	s-
	before high V	ot-	ar-	ha-
	before obstruent	ot-	ar-	ha-
	before other C	oh-	ar-	ha-

Lenakel is the only language to show allomorphic variation in the trial; and since this exactly matches the variation in the plural, I assume that it is the result of a spreading effect. Finally, South-west Tanna alone shows some variation in non-subject number markers: -asəl '3' and -aua 'p' occur after consonants, while -səl '3' and -a 'p' occur after vowels; the variation in the dual, however, does not admit of such a simple explanation, although the form -u marking IIIns may involve dissimilation (expected /ili-lu/ producing iliu).¹⁶

3.2 Dual

The dual non-subjective suffix is reconstructible as *(rL)au, the ambiguity in the initial consonant being due to the fact that the Erromangan languages, which do not mark number distinctions in non-singular non-subjective pronouns, are the only SV languages to reflect the distinction between *r and *L. The forms of the initial consonant of the North Tanna, Whitesands and Lenakel suffixes clearly support the final *i which has been reconstructed for most non-subjective pronominal bases - with no *i in the immediate environment, the suffix -iau would be expected - and thus provides further evidence for the metathesis of POC *kinta Iins as PSV *gadi.

The dual subjective prefix is also, I believe, reconstructible as *(rL)au-. This form is only retained in full as one of the Kwamera allomorphs. The loss of the initial consonant in Anejom has a fairly natural explanation, since the sequences int-rau Ii2, eyr-rau Ix2 and er-rau III2 would fairly naturally condition loss of the r of the suffix, and this would, again naturally, have spread to the form for II2.

The other Tanna languages show two allomorphs: one deriving from *(rL)a- occurs basically before consonants (plus high vowels and, sometimes, o functioning as glides in some environments), while the other deriving from *u- occurs basically before vowels. In an earlier paper (Lynch 1977b:14) I suggested a possible phonetic motivation for this, which is repeated below using Lenakel as the example:

(35)	*(rL)au-kani	*(rL)au-a-toka	POC/PSV form
	2-eat	2-verb:marker-stay	
	[yaw-kani]	[yaw-a-toka]	phonetic equivalent
	[ya-kani]	[w-a-toka]	REDUCTION
	[ya-kən]	[w-a-rək]	other regular rules
	/ia-kən/	/u-arək/	underlying form

Basically, the hypothesis suggests that the [w] (< *u) of the prefix might well be lost before a consonant-initial verb, by a rule of cluster-simplification; while since there would be no motivation for loss of the [w] before a vowel-initial verb, there might well be contraction of the prefix - particularly since a Lenakel verb may admit up to six prefixes before a root. I would thus suggest that the allomorphs in the Tanna languages are derived, by some such process, from *(rL)au-.

Sie shows a specifically dual form only in the first person. However, it seems logical to treat u-, which marks both dual first person and plural non-first person, as the historical dual, and to derive it from *(rL)au- in the same way as the prevocalic Tanna allomorph is derived; presumably, the same contraction which occurred in some environments in Tanna occurred in all environments in Sie.

I thus reconstruct a dual marker $*(rL)au$ which functioned as a prefix to verbs marking number of subject and as a suffix to focal, objective and possessive pronominal bases.

3.3 Trial

In general, the Tanna forms suggest a trial marker $*ha(lrL)i$, to which may have been accreted an initial *a* in the suffixed form in most of these languages once the final vowels of the pronominal bases had been lost.¹⁷ There are, however, a few minor problems which need to be briefly discussed.

Of minor importance are the accretion of initial *r* in the Kwamera non-subjective suffix *-rahar*, possibly due to contamination from the dual; the discrepancy between the vowels of the Lenakel prefixed and suffixed forms (*hal-* and *-hel*); and metathesis in the South-west Tanna subjective prefix. More important, however, is the discrepancy between the consonants in the South-west Tanna prefixed (*lh-*) and suffixed (*-(a)səl*) forms. Although there is some variation between *s* and *h* in Tanna, the direction of the change (which is more pronounced in Erromango) is clearly $s > h$ and not $h > s$: one form with *h* among a number of forms with *s* could be explained as due to the effects of this sound change beginning to operate; the reverse, however, does not admit of such a simple explanation. At this stage, I can only suggest that there has been some contamination from the numeral (*kəsisəl*) in South-west Tanna which has affected the fairly transparent number-suffix, but not the more morphologically embedded verbal prefix.

Since no plural reconstruction to be proposed in 3.4 below could have given rise to the Sie first person plural marker *l-*, I suggest that this derives from the same, or a similar, reconstruction, by the same kind of reduction process by which dual *u-* derives from $*(rL)au$.

Now the Anejom form is *taj*. Ignoring for the moment the initial consonant, the remainder of the form is cognate with the remainder of the Tanna-based reconstruction, and in fact disambiguates the second consonant - Anejom *aj* can only derive from $*ali$. The difficulty arises with the initial consonant: Anejom θ , and not *t*, is cognate with Tanna $*h$.¹⁸ I incline to the view that the Anejom form represents an irregular development from PSV $*hali$. Given the normal development of $*hali$ as θaj , the forms of the trial focal and subjective pronouns would have been:

(36)	Focal	Subjective
Ii3	akat- θaj	int- θaj
Ix3	ajam- θaj	e γ r- θaj
II3	ajou- θaj	ek- θaj
III3	ar- θaj	er- θaj

Now it is not improbable that θ may have assimilated to the non-continuant quality of a preceding stop (*t*, *k*) or flap (*r*) - a situation which would have obtained in six of the eight forms listed above. I propose, then, that Anejom *taj* '3' is an irregular, but explainable, development of a PSV trial marker $*hali$.

3.4 Plural

In discussing the marker of plural, I am ignoring the Erromangan languages, since I treated the first person plural verbal prefix l- as an historical trial, and the second and third person verbal prefix u- as an historical dual.

No clear single reconstruction can be made here. North Tanna and Lenakel forms suggest a reconstruction *at(V) (where *V is not *i); the Whitesands *subjective* marker also derives from this reconstruction. On the other hand, South-west Tanna and Kwamera forms suggest the following possible reconstructions: *a(ha) or *(a)sa, ultimately derivable from a PSV form *(a)tia. The Whitesands *non-subjective* marker would also derive from this reconstruction, although the South-west Tanna non-subjective allomorph -aua remains unexplained. The Anejom form -a is consistent with *both* of the above reconstructions, since the historical rule of final consonant loss is far more thoroughgoing in its application in Anejom than in other SV languages (cf. Lynch 1978a:736-737).

I would suggest that, just as it appears to be necessary to reconstruct doublets for the numeral *four*, so it is probably also necessary to reconstruct doublets for the plural marker, which is phonologically related to the form for *four*. Unfortunately, the absence of reflexes in Erromango and the ambiguity of the Anejom reflex means that *neither* reflex can be attributed with confidence to PSV.

3.5 Development

The relevant reconstructions are summarised below.

(37)	POC	PEO	PSV numerals	PSV number-markers
two/dual	*dua	*dua	*-Lua	*(rL)au
three/trial	*tolu	*tolu	*-s(ie)li	*hali
four/plural	{*pat *pati	- *pati	*-vat(V) *-vati	*at(V) *(a)tia

It will be immediately obvious that, while the PSV numerals unambiguously *derive* from the POC/PEO forms, the PSV number-markers are merely phonologically *related* to those forms.

The dual number-marker, apart from the unfortunate ambiguity in the consonant, shows clear and unpredictable metathesis of the vowel cluster. Although a number of Oceanic languages show reduction of the numeral when it functions as a number-marker, I am aware of only one other area where similar *metathesis* is found: cf. Standard Fijian rua *two*, but keirau Ix2, kemudrau II2, irau III2.

The trial number-marker shows an unpredictable change in both the consonant and the vowel of the first syllable. I am not aware of similar changes in other Oceanic groups.

The two reconstructed plural number-markers show the same development from the numeral - unexpected loss of the initial consonant (PSV *v). A similar development occurs in the New Britain-New Ireland area.

3.6 Subjective number-marking

It is clear from what has gone before that the non-subjective pronouns in PSV consisted of a person-marking base + a number-marking suffix, the two functioning as a unit. What is *not* clear is the way in which markers of both person-of-subject and number-of-subject functioned in the PSV verbal complex.

The Erromangan and Tanna languages show considerable morphological complexity of the verbal unit. The two examples given below would be morphologically 'longer' than the statistical norm, but nevertheless natural utterances:

(38) SIE: *yau-pe-tu-am-anh-i* (> *yaupetumanhi*)
I-conditional-negative-PRES-IRR:*see*-III
I might not see it

LEN: *t-i-ak-am-etu-ar-ol?*
future-Ix-CONC-CONT-interrogative-p-*do*
How will we be doing it?

Only in ('Old') Anejom do the markers of person-of-subject and number-of-subject always occur adjacent to each other. Apart from Kwamera, the other languages of Erromango and Tanna resemble Sie and Lenakel: person-markers may be, and quite often are, separated from number-markers by one or more other prefixes. Kwamera differs in that person-markers are normally immediately followed by number-markers. However, Monty Lindstrom suggests (personal communication) that this may be the result of a fairly recent change, and that there are older forms, still acceptable, which allow tense/aspect markers to intervene between markers of person and number.

It appears, then, that there was no person + number 'pronominal unit' in PSV. Although such a unit could be postulated on the basis of the structure of non-subjective pronouns, it is clear from this analysis of subjective person- and number-marking that each 'component' functioned separately.

4. CONCLUSION

The PSV pronominal system, and the reconstructed forms within that system, are clearly derived, but show considerable departures, from the system and forms reconstructed for POC and PEO. Formal developments have been discussed elsewhere in this paper. I would, however, like to say something about systemic developments.

As it is presently reconstructed, POC shows a singular/plural distinction only, with the plural forms being monomorphemic: i.e. there is no plural-marking morpheme as such. PEO has been reconstructed as having a four-number pronominal system; however, the plural pronouns are monomorphemic, while the dual and trial pronouns consist merely of plural pronoun + unaltered numeral. The PSV system shows two significant departures: the development of a specifically plural suffix, which means that the historical plural pronoun becomes a non-singular bound base form; and the development of number-markers which are not numerals (full or reduced). The full system is retained in Anejom and the Tanna languages, though it is lost in the Erromangan languages, which look more POC-like.

This paper has attempted to reconstruct the PSV system only. It will be of interest for wider subgrouping purposes, however, to examine other subgroups for formal and systemic correlates with the various PSV 'innovations'. That, however, remains a matter for future research.

NOTES

1. Vanuatu was formerly known as the New Hebrides, and the Tafea District as the Southern District. The Southern Vanuatu subgroup was previously referred to as the South-Hebridean subgroup (Lynch 1978a).
2. One of the rules referred to simplifies a cluster of two coronal consonants across a morpheme boundary by deleting the first: /kat-lau/ > kalau Ii2, /il-lau/ > ilau III2; cf. /n-əm-ar-su/ (II-PST-p-plant) > nəmasu you (pl.) *planted (it)*. The other rule metathesises i and h in i-hV clusters: /kami-hel/ > kamhiel III3; cf. /i-hapwu/ (personaliser-smash) > hiapwu (as in hiapwu kopwiel *one who smashes rocks*) = *west wind*.
3. Data for extinct Sorung are restricted to possessive pronouns, while data for Utaha, also extinct, are unreliable. Whitesands forms which differ from those given in Lynch 1974 are corrections based on a 1984 field trip.
4. Backing and rounding of *a to o before *u or *w and subsequent reduction of the ou-cluster to o is not uncommon in the SV languages: e.g. POC *mauRi *left hand* > WSN moul, KWM mour, SIE mwor; *awaŋ *be open* > WSN ouaŋ, NTN oaŋ, etc. (Note: most POC and PEO forms quoted in this paper may be located in Wurm and Wilson 1975.)
5. Anejom e is the unconditioned reflex of POC *i (Lynch 1978a:747): e.g. POC *kita *see* > ANJ e/yeŋ; *pinsa *how much?* > e/heθ; *qanusi *spit* > aŋθe/i; *tapine *woman* > na/taheñ; etc.

It appears necessary to reconstruct two PSV velar stops *k and *g - the latter probably prenasalised - as well as a velar fricative *ɣ. Although there is sporadic variation, including some cases of unexplained loss, we can tentatively establish these protophonemes on the basis of the following patterns:

	*g	*k	*ɣ
Erromango:	URA g, others k	all k	all ɣ
Tanna:	all k	all k	NTN, WSN ŋ, LEN k, Ø, SWT, KWM Ø
Anejom:	k	k, ɣ	ɣ, Ø

PSV *g does not occur frequently, being found mainly in pronouns; cf. the discussion in 2.1.2 and 2.2.2 below.

6. For example, POC *kani *eat* > ANJ ɣiñ; *taŋmane *man* > na/tamwañ; etc.
7. Although *d does not occur with any great frequency, cases of palatalisation of *t before *i are the rule; e.g. POC *kati *bite* > LEN kəs, KWM ahi, ANJ a/yeŋ; POC *puti *banana* > SIE no/voh, SWT nu/kwus, ANJ no/hos; etc.
8. Three liquids need to be reconstructed for PSV: *l, *r and *L. South-west Tanna merges all three liquids as l, and Kwamera merges them as r. North Tanna, Whitesands and Lenakel also merge all three, as l adjacent to *i and as i in other environments. The remaining correspondences are:

	*l	*r	*L
Sie:	d-, -l-, -l	r	r
Ura:	l	r	l
Utaha:	l (?)	r	l-, -r-, -l (?)
Anejom:	j/___*i, l else	Ø/___#, r else	Ø/___#, r else

9. See, for example, in previous notes the reflexes of POC *kita (note 5), *kani (note 6), *kati (note 7), and also, e.g. POC *kutu *louse* > SIE no/ʔut, URA wit, NTN, WSN kə/ŋət, LEN kur, SWT kel, KWM ur, ANJ ne/ʔet.
10. Note that, in reconstructing Proto-Erromangan (Lynch 1983e, esp. pp.195-199), I was unable to find any examples of *any* of the prenasalised stops (*b, *d and *g) occurring finally. It may well be that the distinction between voiced prenasalised and simple voiceless stops was neutralised in that position.
11. There is one other case of Sie final ŋ corresponding to k in other SV languages: SIE paŋ, WSN, LEN, KWM nə/pək, ANJ in/pak *banyan* < POC *mpaka. Unfortunately I have no other Erromangan data on this form. The interpretation being made here would suggest a PSV reconstruction *(pb)aga.
12. This topic is more fully discussed in Lynch 1982b:116-118; cf. also Lynch 1982c.
13. The correspondence NTN, WSN, ANJ t, LEN, KWM r, SWT l is well attested: e.g. POC *tama *father* > NTN, WSN təm-, ANJ e/tma-, LEN rəm-, KWM remu-, SWT ləm-; *mata *eye* > WSN nə/mt-, ANJ nesŋa-ne/mta-, LEN nə/mr-, SWT nəml-, etc.
14. At first glance, the forms URA il-, ANJ er- (along with Utaha e|-) suggest a second PSV reconstruction for IIIns - *il(V)-. Since, however, the focal for IIIns was reconstructed as *ili, it appears that this form is a 'short pronoun' and not a subject prefix.
15. See Lynch 1977b and data listed in Tryon 1976 which suggest that *front* vowel reflexes in both syllables of POC *tolu are quite widespread.
16. A number of minor details, irrelevant to the present discussion, have necessarily been glossed over here. A fuller picture is presented in Lynch 1977b and in the published grammars of individual languages.
17. Note here the synchronic distribution of the non-subjective trial suffixes in South-west Tanna, discussed above in 3.1.2.
18. For example (using Lenakel as a representative Tanna language), POC *kansupe *rat* > LEN kahau, ANJ in/ʔeθo; *masakit *sick* > LEN a/mha, ANJ e/mθa; *qanusi *spit* > LEN aŋh, ANJ aŋθe/i; etc.
19. In the first person plural, the number marker l occurs twice: koplemlaŋhi (= /ko-p-l-am-l-aŋh-i/ Iins-conditional-p-PRES-p-IRR:see-III) *we might see it*.

REFERENCES

- CAPELL, A.
 1960 *Anthropology and linguistics of Futuna-Aniwa, New Hebrides*. Sydney: Oceania Linguistic Monographs No.5.
 n.d. A re-study of the language of Aneityum, New Hebrides. TS.
- CAPELL, A. and John LYNCH
 1983 Sie vocabulary. In Lynch, ed. 1983:75-147.

CLARK, Ross

- 1985 Languages of north and central Vanuatu: groups, chains, clusters and waves. In Andrew Pawley and Lois Carrington, eds *Austronesian Linguistics at the 15th Pacific Science Congress*, 199-236. PL, C-88.
- n.d. Some proto-forms from north and central Vanuatu with selected supporting evidence. MS. Mimeographed.

GERAGHTY, Paul A.

- 1983 *The history of the Fijian languages*. Honolulu: Oceanic Linguistics Special Publication No.19.

INGLIS, J.

- 1882 *A dictionary of the Aneityumese language*. London: Williams and Norgate.

LINDSTROM, Lamont

- 1984 Kwamera dictionary - N̄ikukua sei nagkiarrien N̄ininife. Computer printout. To be published as PL, B-95.

LYNCH, John

- 1974 Some wordlists from the southern New Hebrides. MS. Mimeographed.
- 1975 *Lenakel phonology*. *Working Papers in Linguistics, University of Hawaii*, 7/1.
- 1977a *Lenakel dictionary*. PL, C-55.
- 1977b On the history of the Tanna numerals and number-markers. *Te Reo* 20:3-28.
- 1978a Proto-South Hebridean and Proto-Oceanic. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*. PL, C-61:717-779.
- 1978b *A grammar of Lenakel*. PL, B-55.
- 1982a South-west Tanna grammar outline and vocabulary. PL, A-64:1-91.
- 1982b Anejom grammar sketch. PL, A-64:93-154.
- 1982c A century of linguistic change in Anejom. MS. Mimeographed.
- 1982d Towards a theory of the origin of the Oceanic possessive constructions. In Amran Halim, Lois Carrington and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, vol.1: *Currents in Oceanic*. PL, C-74:243-268.
- 1983a The languages of Erromango. In Lynch, ed. 1983:1-10.
- 1983b Ura grammar sketch and vocabulary. In Lynch, ed. 1983:148-183.
- 1983c Utaha. In Lynch, ed. 1983:184-190.
- 1983d Sorung. In Lynch, ed. 1983:191-192.
- 1983e Preliminary remarks on Proto-Erromangan. In Lynch, ed. 1983:193-220.

LYNCH, John, ed.

- 1983 *Studies in the languages of Erromango*. PL, C-79.

LYNCH, John and A. CAPELL

- 1983 Sie grammar outline. In Lynch, ed. 1983:11-74.

LYNCH, John and D.T. TRYON

- 1985 Central Oceanic: a subgrouping hypothesis. In Andrew Pawley and Lois Carrington, eds *Austronesian Linguistics at the 15th Pacific Science Congress*, 31-52. PL, C-88.

PAWLEY, Andrew

- 1972 On the internal relationships of Eastern Oceanic languages. In R.C. Green and M. Kelly, eds *Studies in Oceanic culture history*, vol.3. Pacific Anthropological Records No.13:1-142. Honolulu: Bishop Museum Press.

ROSS, Malcolm

- 1981 Proto-Oceanic from the top down. UPNG: mimeographed.
- 1982 The development of the verb phrase in Oceanic languages of the Bougainville region. In Amran Halim, Lois Carrington and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, vol.1: *Currents in Oceanic*. PL, C-74:1-57.
- 1983 Proto-Central Papuan morphosyntax. ANU: computer printout.

TRYON, D.T.

- 1976 *New Hebrides languages: an internal classification*. PL, C-50.

WALSH, D.S.

- 1982 The restricted distribution of Raga /k/: starting point for a subgrouping hypothesis. In Rainer Carle et al., eds *Gava^c: Studies in Austronesian languages and cultures, dedicated to Hans Kähler*, 237-251. Berlin: Dietrich Reimer.

WURM, S.A. and B. WILSON

- 1975 *English finderlist of reconstructions in Austronesian languages (post-Brandstetter)*. PL, C-33.

THE SOUND SYSTEM OF PROTO-CENTRAL-PACIFIC

Paul Geraghty

1.1 INTRODUCTION

The theory that the languages of Fiji, Rotuma, and Polynesia form a closed subgroup was first proposed by Grace (1959). He later (1967) named the subgroup "Central Pacific", and the name has become generally accepted.

Blust (1976), Pawley (1972, 1979), Geraghty and Pawley (1981), and Wilson (1982), among others, have assumed the Central Pacific (CP) hypothesis, and some Proto-Central-Pacific (PCP) lexical items have been reconstructed in Blust 1976 and Geraghty and Pawley 1981.¹ However, as I have argued (Geraghty 1983:352-366), a compelling case for Central Pacific has yet to be made, all of the innovations claimed by Grace (1959) and Pawley (1972) to characterise PCP being either shared only by Polynesia and parts of eastern Fiji, or invalid in some other way. It is not my intention here to discuss further the validity of the CP subgroup, but to provide a firm basis for further discussion by attempting to reconstruct the sound system of PCP, and outlining its development in the daughter languages. No internal subgrouping is as yet assumed, so forms witnessed in two of the three major witnesses, or in any of these plus an external witness, are reconstructed.² This reconstruction is largely based on proposals made in Geraghty 1983 with respect to Proto-Eastern-Oceanic, with one additional phoneme, some phonetic and orthographic modifications, and considerable additional data, especially from Rotuman.

1.2 ORTHOGRAPHY AND SOURCES

Unless otherwise indicated, phonetic values in all data and reconstructions in this paper, regardless of source, are as follows: a,e,f,h,i,k,l,m,n,ñ,o,p,r,s,t,u,w,y,z,ʔ as written; b[mb],c[ɔ̃], d[nd], dr[ndr], g[ŋ], j[tʃ], q[ŋg], v[v] or [β], x[x]; vowel length is indicated by a macron.

All glosses are written according to the conventions described in Geraghty 1983:8-13.

In choosing symbols for reconstructed PCP phonemes, I have been guided by two major considerations: phonetic suitability (but with preference for single letters of the Roman alphabet over digraphs and exotic phonetic symbols), and orthographic usage in daughter languages.

Fijian data are from my fieldnotes, and written in the orthography described in Geraghty 1983:4-8. Proto-Fijian reconstructions are likewise my own.³ Note

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 289-312. *Pacific Linguistics*, C-94, 1986.

© Paul Geraghty

that one of the major differences between Proto-Fijian and Standard Fijian ('Bauan') is that Standard Fijian (SF) has undergone Eastern Fijian Apical Prenasalisation (Geraghty 1983:74-96), resulting in SF *d*, *dr*, and *s* from PFJ **t*, **r*, and **c*, respectively, in initial position in many common nouns.⁴

Proto-Polynesian (PPN) data are mostly from Biggs 1978, 1979, Ranby 1980, and Geraghty 1983.

Occasional reference is made to the following external witnesses: Proto-Southeast-Solomons (PSS) (Levy n.d.), Proto-Micronesian (PMC) (Bender et al n.d.), and Proto-North-Central-Vanuatu (PNCV) (based on data in Guy 1978, Clark 1985, and Walsh 1984).

Rotuman data are from Churchward 1940, with some additional definitions from informants. Diacritics for umlaut, which is predictable in all citation forms, are omitted. Unless otherwise stated, Rotuman forms cited are "directly inherited" (Biggs 1965), that is, not Polynesian loans. (PN?) after a form means that, on purely phonological grounds, it may be a Polynesian loan.

2.1 THE SOUND SYSTEM

The sound system proposed is shown in Table 1.⁵

Table 1: The PCP sound system

	bilabial		alveolar		palatal	velar	labio-velar	glottal
		dental	liquids	fricatives				
fricatives	v			c	z	x		
stops	p	t	r			k	kw	ʔ
prenasalised obstruents	b	d	dr	s	j	q	qw	
nasals	m	n	l		ɲ	g	gw	
glides	w				y			

Note that the table has been compressed somewhat, so the place and mode of articulation labels are not necessarily to be interpreted strictly. For example, it is not claimed that **s* was phonetically prenasalised, or that **l* was a nasal.

In the following sections, we will examine the system by place of articulation, discussing phonetic values and reflexes. Examples will, as far as data permit, illustrate reflexes of consonants in both initial and medial position, and before front, low, and back vowels.

2.2 BILABIALS

Table 2: The reflexes of the PCP bilabials

PCP	v	p	b	m	w
PFJ	v	p	b	m	w
PPN	f	p	p	m,∅	w
ROT	h,∅/?	p	p	m,∅	v

Examples:

v-

- **vitu seven*: PFJ **vitu*, PPN **fitu*, ROT *hifu*
 **vanua land*: PFJ **vanua*, PPN **fanua*, ROT *hanua*
 **vutu k tree, Barringtonia asiatica*: PFJ **vutu*, PPN **futu*, ROT *hufu*

-v-

- **avi fire*: PFJ **yavu burn*, PPN **afi*, ROT *rahi*
 **tuva k vine, Derris trifoliata*: PFJ **tuva*, ROT *fuha*
 **mava heavy*: PPN **mamafa*, ROT *maha*
 **tavu set fire*: PFJ **tavu*, PPN **tafu*, ROT *fahu*

p-

- **pisi-k squirt*: PFJ **pisi-k*, PPN **pisi-k*, ROT *pusi burst, splash*
 **pā trolling hook*: PFJ **pā*, PPN **pā*
 **popo (wood) rotten*: PFJ **popo*, PPN **popo*, ROT *popo* (PN?)

-p-

- **ripi sharp edge*: PFJ **ripi shin*, PPN **lipi*
 **sape (foot) malformed*: PFJ **sape*, PPN **sape*, ROT *tape* (for **sape*) *kick w toe*

b-

- **bebe butterfly, moth*: PFJ **bēbē*,⁶ PPN **pepe*, ROT *pepe* (PN?)
 **bā wall, fence*: PFJ **bā*, PPN **pā*, ROT *pā* (PN?)
 **buto- navel*: PFJ **buto-*, PPN **pito*, ROT *pufa*

-b-

- **kabe string*: PFJ **kabe string from coconut stem*, ROT ?*ape*
 **tubā k land crab, Cardisoma*: PFJ **tubā*, PPN **tupa*,⁷ ROT *fupa*⁷
 **tubu grow*: PFJ **tubu*, PPN **tupu*, ROT *fupu*

m-

- **miji suck*: PFJ **misi*, PPN **miti*
 **mata- eye, face*: PFJ **mata-*, PPN **mata*, ROT *mafa*
 **moze sleep*: PFJ **moze*, PPN **mohe*, ROT *mose*

-m-

- **kumete bowl*: PFJ **kumete*, PPN **kumete*, ROT ?*umefe*
 **cama outrigger float*: PFJ **cama*, PPN **hama*, ROT *sama*
 **ñamu mosquito*: PFJ **ñamu*, PPN **namu*, ROT *ramu*

w-

- *wī *k tree*, *Spondias dulcis*: PFJ *wī, PPN *wī, ROT vī (PN?)
 *weka *bird*, *Rallus*: PPN *weka, ROT ve?a
 *waqa *canoe*: PFJ *waqa, PPN *waka, ROT vaka (PN?)

-w-

- *kauki *sand crab*: PFJ *kauki, PPN *kawiki, ROT ?avi?i
 *kawi *fish-hook*: ROT ?avi (POC *kawil)
 *tawa *k tree*, *Pometia pinnata*: PFJ *tawa, PPN *tawa, ROT fava

The reasons for reconstructing *v rather than *f are not strong, simply that *v is a more common cognate in external witnesses (PSS, PNCV) than *f (PMC). The distinction between PCP *b and *p is maintained only in PFJ, and only on evidence from parts of eastern Fiji, but is supported by evidence from the Solomons (Geraghty 1983:103-114).

Some instances of PCP *v become ∅ or ? in Rotuman:

INITIAL

- *vaka-V causative > a?a
 *vu(cz)u *box*, *punch* > ?usu
 *vusi *tie in a bunch* > usi
 *vu?u- *tree* > u- prefix to some tree names

MEDIAL

- *V(cz)ivo *down* > sio
 *tovu *sugarcane* > fo?u
 *uvi *blow* > ui
 *vavine *woman* > haina

It is probably significant that the most common environment is before a high back vowel, with two before a high front vowel, and only one each before o and a. Both cases of glottal stop are before u.

Perhaps related to this change is the sporadic loss of intervocalic *m before *u, which occurs in both PPN and Rotuman:

- *kamu Iip > ROT ?au (Hale 1846:472 also recorded ?amu)
 *malumu *soft* > PPN *malū
 *N-mu III > PPN *-u, ROT -u

The same change occurs sporadically in Waidina, eastern Vitilevu (Geraghty 1983: 178-179).

2.3 DENTALS

Table 3: The reflexes of the PCP dentals

PCP	t	d	n
PFJ	t	d	n
PPN	t	t	n
ROT	f/j/s	t/j	n

Examples:

t-

*tinaʔe *intestines*: PPN *tinaʔe, ROT finae
 *taliga- *ear*: PFJ *taliga-, PPN *taliga, ROT faliga
 *tuna *Anguillidae, freshwater eel*: PFJ *tuna, PPN *tuna, ROT funa

-t-

*ʔoti *finished*: PFJ *oti, PPN *ʔoti, ROT ofi
 *mataʔu *right-hand*: PFJ *mataʔu, PPN *mataʔu, ROT mafau
 *ʔatu *line, row*: PFJ *yatu, PPN *ʔatu, ROT afu

d-

*degu *nod*: PFJ *deguvacu *raise eyebrows in assent (vacu eyebrow)*, ROT tegi
 *dagwa *loose, slack*: PFJ *dagwa, PPN *tagataga
 *dañudañu *fallow*: PFJ *da(nñ)uda(nñ)u, ROT taitai (POC, PPN *talu)
 *dui *different*: PFJ *duidui, ROT tū

-d-

*vidi *spring*: PFJ *vidi, PPN *fiti, ROT hiti *start w surprise*
 *voda *rocks in sea*: PFJ *voda, PPN *fota
 *mudu *cut off, sever*: PFJ *mudu, PPN *mutu, ROT mutu (PN?)

n-

*niu *coconut*: PFJ *niu, PPN *niu, ROT niu (PN?)
 *na(czs)u *roast, bake*: ROT nasu
 *natu *mash, knead*: PFJ *natu, PPN *natu
 *novo *sit, stay*: PFJ *novo *lie still*, PPN *nofo, ROT noho

-n-

*kini *pinch*: PFJ *kini, PPN *kini, ROT ʔini
 *kanace *k fish, Mugil, mullet*: PFJ *kanace, PPN *kanahe, ROT ʔanasi (for *ʔanase)
 *tunu *cook*: PFJ *tunu *reheat (food)*, PPN *tunu *cook on open fire*, ROT funu *cook by boiling*

The j and s reflexes in Rotuman are somewhat problematic. It appears that *t, before shifting to f (via an intermediate stage [θ], recorded by Hale (1846) and Turner (1884)), assimilated to a following j or s (from *c, *s, or *z):⁸

*ta(cz)i *sea* > sasi
 *tazi-ña *his/her younger sibling* > sasiga

*tali(cz) *e k tree, Terminalia catappa* > salisa *k edible almond-shaped fruit*
 *taji *shave* > jaji

There are no counter-examples in my data. The same rule applies to *d in the one eligible form:

*du(cj)(iu) *point* > juju

and to *s before *j:

*sije *k fish, Hemirhamphus, garfish* > jija

There are, however, two further cases of *d becoming Rotuman j where assimilation does not appear to be involved:

*dulɿ *k bird, plover* > juli

*donu *right, correct* > nojo (metathesis)

Two hypotheses suggest themselves. That j is the regular reflex of *d before *u is, however, contradicted by tū *different* (< *dui) and tutuʔu *k fish, small, black* (< *duku *k fish, Abudedefduf sp*). It is more likely that j reflects *d before l or n. Although not a particularly plausible environment, the only apparent counter-example, tulou *millipede* (< *dolou *earthworm*) may have been ineligible for the change because of stress placement, or may be a loan from an unknown source. The problem, of course, requires more data.

2.4 LIQUIDS

Table 4: The reflexes of the PCP liquids

PCP	r	dr	l
PFJ	r	dr	l
PPN	r,l	r,l	l
ROT	r/∅	t	l

Examples:

r-

*riki *small*: PFJ *riki, PPN *riki, ROT ririʔi (plural)

*rano *lake*: PFJ *(rdr)ano, PPN *rano +*swamp*, ROT rano *swamp*

*rua *two*: PFJ *rua, PPN *rua, ROT rua

-r-

*iri *fan*: PFJ *iri, PPN *iri, ROT iri (PN?)

*viri *plait*: PFJ *viri *lash (fence, raft+)*, PPN *firi, ROT hiri

*mara *fermented food*: PFJ *mara *stench*, PPN *mara, ROT mara (Hale 1846) (PN?)

*curu *enter, go through*: PFJ *curu, PPN *huru, ROT suru

dr-

- *driudriu *k small ant*: PFJ *driudriu, ROT tuitui (metathesis)
 *dram(iu) *chew*: PFJ *dram(iu) *lick*, PPN *lam(iu), ROT tami
 *dranu *bathe in fresh water*: *dranu, PPN *ranu
 *drumani *k edible sea-anemone*: PFJ *dr(ou)mani, PPN *rumane, ROT nunami (metathesis and assimilation)

-dr-

- *vadra *Pandanus*: PFJ *vadra, PPN *fara, ROT hata
 *madra *cooked, fermented*: PFJ *madra, PPN *mara
 *tadruku *Chiton*: PFJ *tadruku (PSS *tadux(iu))

l-

- *lima *five*: PFJ *lima, PPN *lima, ROT lima
 *lago *k insect, fly*: PFJ *lago, PPN *lago, ROT laga
 *lua *vomit*: PFJ *lu(ae), PPN *lua, ROT lua *spit*

-l-

- *taliga- *ear*: PFJ *taliga-, PPN *taliga, ROT faliga
 *zala *path, road*: PFJ *zala, PPN *hala, ROT sala
 *walu *eight*: PFJ *walu, PPN *walu, ROT valu

PCP *r apparently becomes Rotuman Ø between high vowels, though there is some contradictory evidence:

- *buru *present food*: PFJ *buru(a), PPN *pulu(a), ROT pū
 *puru- *abdomen, thorax*: PFJ *poro-, ROT pū (PMC *pur(iu)a *aesophagus, gullet, stomach*)
 *riri *shed, hut*: PFJ *riri, ROT rī *house*
 *tiri: PPN *ti(ri)i (*woman*) *fertile*, ROT fī (*woman*) *prolific*
 *turu- *knee*: PFJ *turu-, PPN *turu, ROT fū
 *xuru *rumble*: PFJ *kuru, PPN *ʔulu, ROT ʔū *bang*

The contradictory data are:

- *curu *enter, go through*: PFJ *curu, PPN *huru, ROT suru
 *viri *plait*: PFJ *viri *lash (fence, raft+)*, PPN *firi, ROT hiri
 *vuvuru *catch (fish, animal) w hand*: PFJ *buburu,⁹ ROT huhuru

There is no evidence that PCP *r and *dr remained distinct in PPN; *dr, like *d and *q, merged with its non-prenasalised counterpart. The resultant *r merged partially with *l, under conditions yet to be determined. The merger was completed in Proto-Nuclear-Polynesian, but not in Proto-Tongic, where *r became Ø. Data available offer some suggestions as to conditions for the merger, but as yet no clear pattern can be discerned.

Examples of PCP *r and *dr > PPN *l:

r-

- *riri *boil*: PFJ *riri, PPN *lili
 *rau- *leaf*: PFJ *rau-, PPN *lau, ROT rau
 *rogo *quiet, silent*: PFJ *rorogo, PPN *logo

-r-

*marari *k fish, wrasse*: PPN *malali, ROT marari (PN?) (PMC *merari)
 *gara *scream, howl*: PFJ *gara, PPN *gala
 *turu *drip*: PFJ *turu, PPN *tulu

dr-

*dreu *ripe*: PFJ *dreu, PPN *leu, ROT toutou
 *dranu *fresh water*: PFJ *dranu, PPN *lanu, ROT tanu *water*
 *druma *shy*: PFJ *druma, PPN *luma

-dr-

*modri *smooth, hairless*: PFJ *modri, PPN *molemole
 *(cz)(eo)dra *asthma*: PFJ *(cz)odra, PPN *sela

2.5 ALVEOLAR FRICATIVES

Table 5: The reflexes of the PCP alveolar fricatives

PCP	c	s
PFJ	c	s
PPN	h,s	s
ROT	s	s/j

Examples:

c-

*cina *illuminate, fish by torchlight*: PFJ *cina, PPN *hina, ROT sina
 *cakau *coral reef*: PFJ *cakau, PPN *hakau, ROT sa?au *rocks and coral on sea bottom*
 *cucu- *breast*: PFJ *cucu-, PPN *huhu, ROT susu

-c-

*kanace *k fish, Mugil, mullet*: PFJ *kanace, PPN *kanahe, ROT ?anasi
 *vuca *rotten*: PFJ *vuca, ROT husa *pus*
 *cucu- *breast*: PFJ *cucu-, PPN *huhu, ROT susu
 *vacu- *eyebrow*: PFJ *vacu-, ROT hasu

s-

*sikwa *net-needle*: PFJ *sikwa, PPN *sika, ROT si?a
 *saga: PFJ *saga *attempt, work on*, PPN *saga *work, make, do*, ROT saga *act quickly* (PN?)
 *sua *scull*: PFJ *sua, PPN *sua, ROT sua (PN?)

-s-

*asi *k tree, Santalum, sandalwood*: PFJ *yasi, PPN *asi
 *wasa *open sea*: PFJ *wasa, PPN *wasa, ROT vasa *far out at sea*
 *lasu: PFJ *lasu *false, tell lie*, PPN *lasu *trick, deceive*

PCP *s, like *t and *d, assimilates to a following *j in Rotuman:

*sije *k fish, Hemirhamphus, garfish* > jija

As with PCP *r > PPN *r,l, there is no obvious conditioning for PCP *c > PPN *h,s - the partial merger of PCP *c with *s in PPN. PFJ *c is taken as a true witness to PCP *c because of its close correlation with cognates in the Southeast Solomons (see Geraghty 1983:130-148).

Examples of PCP *c > PPN *s:

c-

*cici *k edible mollusc, inc. Neritidae*; PFJ *cici, PPN *sisi, ROT sisi

*cakule *search for lice*: PFJ *cakule, PPN *sakule, ROT saʔule

*cunu: PFJ *cunu *steam (st)*, PPN *sunu *singe*, ROT sunu *hot*

-c-

*cici *scoop out, gouge out, peel*: PFJ *cici, PPN *sisi, ROT sisi *peel, strip off (skin)*

*macaki *illness, disease*: PFJ *macake *k disease, thrush*, PPN *masaki, ROT masaʔi

Note that PCP *c is considered to be the non-prenasalised member of the *c-s pair. The reasons are that *c occurs as the final consonant in PCP bases while *s, like the phonetically prenasalised obstruents, does not; and that when East Fijian Apical Prenasalisation occurred, *c became s under exactly the same conditions that *t and *r became phonetically prenasalised (Geraghty 1983:90-95). This view was in fact held, for Fijian, by Dempwolff (1934-1938:II:138), but later reversed, apparently by Elbert (1953), followed by Biggs (1965:385) and Pawley (1972:27), presumably for phonetic reasons, the voiced member taken to be more likely to reflect a prenasalised obstruent. The old position here re-affirmed has more recently been argued for by Milke (1961), Hockett (1976:191-192), and Haudricourt and Ozanne-Rivierre (1982:31).

2.6.1 Palatals

Table 6: The reflexes of the PCP palatals

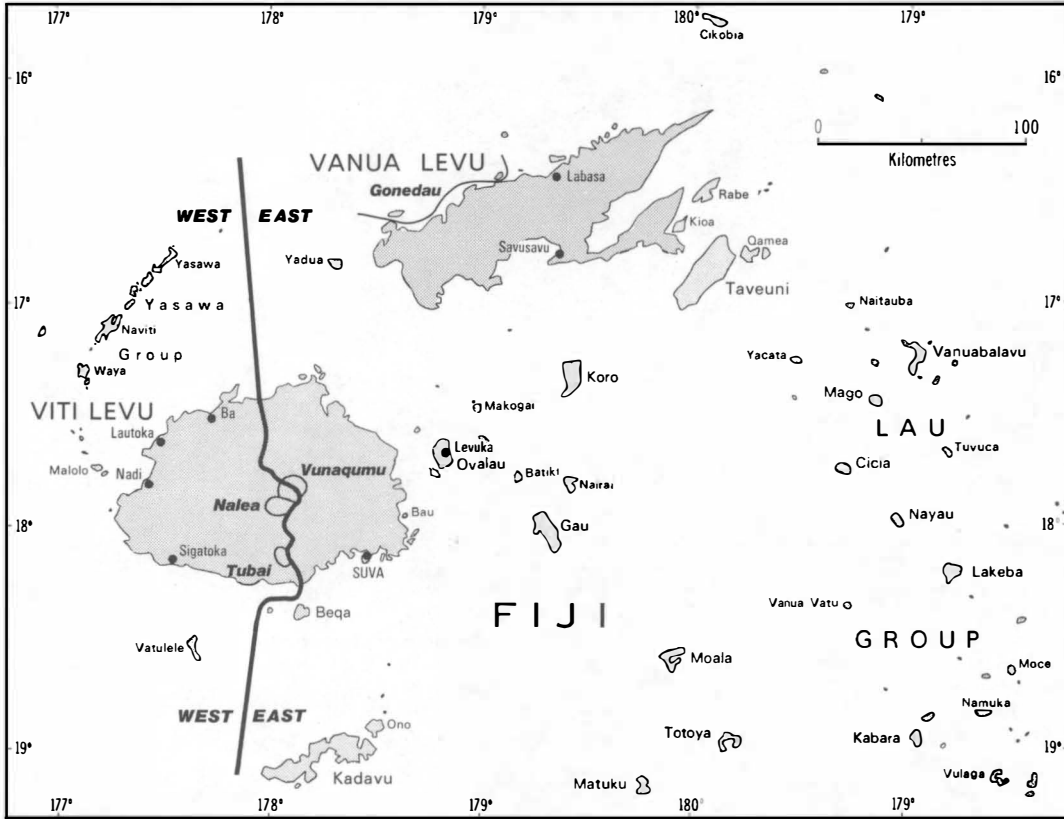
PCP	z	j	ñ	y
PFJ	z	s	ñ/n	c
PPN	h,s	t,s	n	∅
ROT	s	j	∅/r,g,n	∅/r

Only in PFJ is *z distinguished from *c, since both yield *h,s in PPN and s in Rotuman. The evidence for PFJ *z was first presented in Geraghty 1983:125,126, 153-155, where it was tentatively labelled *C. In most Fijian communalects it is regularly realised as c, which is also the reflex of PCP *c. Where *z differs from *c is that in four communalects, two belonging to the Western subgroup, two to the Eastern, it is realised as s (or y/∅), not c. The two Western communalects, Nalea and Tubai, are historically closely related, but now separated.

Tubai is not a totally reliable witness, containing many loans from both Eastern and Western communalects acquired during the prehistoric wanderings of its speakers. One of the Eastern communalects, Vunaqumu, is contiguous to Nalea, but is considered an independent witness because it belongs to a different first-order subgroup of Fijian. Data from Vunaqumu are, however, sketchy, because the last speaker died about 50 years ago, and the data have been culled from the memories of old people who heard it in their youth. The fourth witness is Gonedau, spoken on the islands of Yaqaga, Galoa, Tavea, and Macuataiwai, off the northern coast of Vanualevu. Map 1 shows the locations of these witnesses to PCP *z, and the evidence is presented in Table 7. Reconstructions based only on y in Vunaqumu or Tubai and c elsewhere are not very secure, since y is also a fairly common sporadic reflex of PFJ *c (Geraghty 1983:126-128). Since PCP *z becomes either c or s, but can hardly have been either, [z] seems to be a reasonable guess at its phonetic nature.

There are five instances of PCP *z becoming PPN *s, rather than the usual *h: PPN *tagi-s *cry*, *fusi *irrigated taro bed*, *kese (doublet *kehe) *different*, *masa (doublet *maha) *dry*, and *sole (for *sola) *carry on shoulder*. It is not possible to tell whether or not the PPN *s reflex of *z occurs under similar conditions to the *s reflex of *c.

PCP *j is the reflex of PEO *j as proposed in Geraghty 1983:149-153 on the basis of the correspondence: Fiji s, PPN *t or *s, Rotuman *j, PSS *d.¹⁰ It approximates to the POC *nj proposed by Milke (1968), and the PCP *c of Blust (1976). Only in Rotuman is it retained as a distinct phoneme. In PSS it merges with the reflexes of *d and *dr, in PFJ with *s, and in PPN with *s and *t or *d. Given this pattern of mergers, it seems likely that *j was the 'prenasalised' counterpart of *z. PCP *j was probably [tʃ] or [ts], like its only unique reflex, Rotuman j (Churchward 1940:13,83). The evidence for PCP *j is presented in Table 8. A number of items included in Geraghty 1983 only on the strength of external evidence, usually PSS *d, are omitted here. As with PCP *r, *c, and *z, the PPN split reflex is problematic. There is no obvious conditioning, only a tendency to *t before back vowels and *s before front vowels.



Map 1: Fiji

showing the East-West language division,
and the approximate location of the com-
munalects crucial to the reconstruction
of PFJ *z

Table 7: Evidence for PFJ *z

Forms that are only reconstructable to Proto-Western-Fijian or Proto-Eastern-Fijian are marked (W) and (E), respectively. Forms in brackets are presumed to be borrowed. A dash means the form with that meaning is not cognate.

PFJ	Nalea	Tubai	other Western	Vunaqumu	Gonedau	other Eastern
INITIAL						
<i>zava what</i>	sava	yava	cava	yava	sava	cava
<i>zei who</i>	sei	(cei)	cei	yei	sei	cei
<i>ziqi divide (food) w fingers (W)</i>	siqi	siqi	ciqi	-	-	-
<i>zālevu path, road¹</i>	sālevu	yālevu	cālevu	sālemu	sālevu	sālevu ²
<i>zola live (W)</i>	solo	-	col(ao)	-	-	-
<i>zavu pronounce, men- tion</i>	savu	yavu	cavu	savu	(cavu)	cavu
<i>zō call (W)</i>	sō	-	cō	-	-	-
<i>zola carry on soulder</i>	-	soya	-	sola	-	cola
<i>zai copulate (E)</i>	-	-	-	yai	-	cai
<i>zaka do, make (E)</i>	-	-	-	yaka	-	caka
MEDIAL						
<i>moze sleep</i>	mose	mose	moce	-	mose	moce
<i>tazi- younger same-sex sibling</i>	tasi-	tai-	taci-	tai-	tasi-	taci-
<i>viza how many</i>	visa	(vica)	vica	visa	visa	vica
<i>yaza- name</i>	yasa	(yaca)	yaca	ya-	yasa	yaca-
<i>buzobuzo white</i>	busobuso	buyobuyo	bucobuco	buyobuyo	-	-bucu
<i>la(zy)a sail³</i>	-	laya	la(cy)a	-	lasa	laca
<i>maza empty of liquid, (tide) low</i>	masa	masa	maca	masa	(maca)	maca
<i>māmaza dry</i>	māmasa	masamasa	macamaca	-	(māmaca)	māmaca
<i>naiza when</i>	-	-	-	-	nesa	naica
<i>uza when</i>	-	-	-	-	usa	uca
<i>Vkeze only, alone</i>	-	-	-	-	kese	kece
<i>kuza how</i>	-	-	ku(cy)a	-	kuse-	kuca
<i>vuzi irrigated taro bed</i>	vusi	(vuci)	vuci	vusi	-	vuci
<i>maziV again, adversa- tive</i>	masi	-	maci	-	-	maci
<i>Vwaza only, merely</i>	-	wasā	-	-	-	waca
FINAL						
<i>lua-z vomit - on</i>	lua-s	lua-∅	lua-c	-	lua-s	lua-c
<i>mT-z urinate - on</i>	mT-s	mT-∅	mT-c	-	mimi-s	mT-c
<i>veka-z defecate - on</i>	veka-s	veka-∅	veke-c	-	veka-s	veka-c
<i>tagi-z cry - for</i>	tagi-s	tagi-∅	tagi-c	-	tagi-s	tagi-c
<i>wavu-z run - for</i>	wavu-s	-	wavu-c	-	ovu-s	-
<i>liga-z see</i>	liga-s	-	liga-c	-	-	liga-c
<i>b(ou)i-z smell</i>	bui-s	bui-∅	bui-c	-	-	boi-c
NOTES						
1. From *zala + *levu <i>big</i> (Vunaqumu lemu <i>big</i>).						
2. s from earlier c by East Fijian Apical Prenasalisation (Geraghty 1983:90-95).						
3. Other than Gonedau, all evidence points to *laya, so the Gonedau form is probably irregular.						

Table 8: Evidence for PCP *j

PCP	PFJ	PPN	ROT
INITIAL			
jamu(?)a (palm) fruit stem	sāmoa	taume spathe (met.)	jamu?a
jamu scraps of food	sabusabu	samu	jamujamu
jao spear	sā	tao	jao
jara slip, slide	sara	tala put on (clothes)	jara
jau strike, beat	sau +tattoo	tatau tattoo	jau
je(?) (ei) k insect	-	se(?)e locust	jei cricket
jei tear, rip	sei	(sae)	jei
jevu splash water	sevu	-	jehu drizzle
je(?)a k bird, Lalage sp	sea	(hs)e(?)a	jea
jexejexe k fish, Arothron spp.	sekeseke	te?ete?e	-
jī k plant, Dracaena	-	tī	jī
jiko- kingfisher	sikorere Artamus, woodswallow	tikotara	-
jila look sideways, squint	-	sila	jila
jila (canoe) sheet	sila	tila	-
jiji slip	sisi	-	jiji creep, crawl
joli pick, gather	-	tolī	joli
jona yaws, (octopus) sucker	sona	tona	jona
jopu nod	sopu	-	jopu
jou (sea) rough	sou splash	sou	jou ripple
jo(bp)u dive	-	sopu	jopu
jove k shellfish	sove barnacle	tofe	
MEDIAL			
baja close together	basa meet, opposite	-	paja
duji point	du(cs)i	tus(iu)	jūju
guju- mouth	gusu-	gutu	nuju
gwajala k fish, Epinephelus	kasala (*gwasala)	gatala	vajala
kaja- (kava) stem	kasa-	kata	-
ikajo, kiajo outrigger boom	ikaso	kiato	-
keju- back of head	kesu-	-	?eju
laja tame	lasa	lata	-
laje coral	lase	lase	laje
majaga- (road+) fork	basaga-	māsaga twin	majaga
maja(?)u clever, expert	-	mata(?)u	majau
miji suck	misi	miti +lick	-
muju cut off	musu	mutu	-
sije Hemirhamphus, garfish	sise	(ise)	jija
taji shave	tasi	tasi	jaji
uja transport, carry (cargo)	usa	uta	-
xujim(ai), xumij(ai) crave fish or seafood	kusima	?umiti	-

2.6.2 The source of PCP *z and *j

It was suggested in Geraghty 1983:154-155 that PCP *z may reflect PAN (Proto-Austronesian) *j, notwithstanding certain irregularities, there being some support in the fact that POC (Proto-Oceanic) did distinguish *j (Blust 1978). No PAN source was suggested for PCP *j. With the increase in data, we are now in a better position to look into the PAN source of both PCP *z and *j, along with the other PCP phonemes that derive from the PAN palatal obstruents.

The following list shows the PAN sources for all PCP items with unequivocal *c, *s, *z, and *j. I do not distinguish here between PAN and PMP (Proto-Malayo-Polynesian), and some final consonants have been omitted or simplified.

PCP *c

aca *rub, grate* < *Sasaq *sharpen (blade)*
 cabo *hold in hand* < *saŋpe
 caga *span* < *zaŋan (or *saŋa *bifurcation*)
 cake *climb* < *sakay
 cala *wrong, err* < *salaq
 cama *outrigger float* < *(cs)a(R)man
 cavu-t *pull up, uproot, pull out* < *cabut
 i/cawa/i *parent-in-law* < *qa(cs)awa *spouse*
 caʔa-t *bad* < *zaqat
 cici *k edible mollusc, inc. Neritidae* < *sisi
 ci(kq)o-v *catch w hands* < *cikep
 cila (*sun+*)*shine* < *silak, *cilak*
 cina *torch* < *sinaR *ray of light*
 ciwa *nine* < *siwa
 i/coka *house-beam* < *se(ŋ)kaŋ *crossbar*
 cucu- *breast* < *susu
 cula *sew, pierce* < *sulam *prick, pierce*
 culi- (*taro, banana*) *sucker* < *suliq
 (g)icu *nose* < *ijun (or *ŋusu *upper lip*)
 kanace *k fish, Mugil, mullet* < *kanasay
 mac(eo)ru *hiccough* < *se(d)u
 toci *cut (leaf) into strips* < *testes *tear up*
 v(iu)cov(iu)co *navel, umbilical cord* < *pusej
 voce *paddle* < *be(R)(cs)ay

PCP *s

(vb)oso *squeeze in hand* < *becel
 lasu *tell lie, deceive* < *la(n)cu
 los(ei) *squeeze, wring out* < *leciŋ *squeeze out, squirt out*
 pisi-k *squirt* < *picik *splash, spray, sprinkle*
 saba-k *slap* < *ca(m)pak
 saqa-t *oppose, crash into* < *ca(ŋ)kaŋ *contradict, oppose*
 saqu-m *snatch* < *ca(ŋ)kem *grasp*
 si(bp)a *cut into strips* < *si(ŋ)pak *split*
 somo *mud* < *cemeD *impure*
 sova *pour, dump* ? < *sebar *scatter about*
 sulu *put on clothes* < *(cs)ulu
 vuso *foam* < *buseq

PCP *z

(?)aza- *name* < *ajan
 maza *dry (tide)low* < *maja
 moze *sleep* ?< *pezem *close eyes, sleep*
 (nñ)a(?)iza *when* < *qizan
 tagi-z *cry - for* < *taŋis
 tazi- *younger same-sex sibling* < *tV-Sua(n)ji
 viza *how much* < *pija
 z(ae)i *who* < *(cs)ai
 zala *path, road* < *Zalan
 zava *what* < *apa, *sapa
 zaʔi-t *copulate* ?< *zaqit *sew, join*
 ʔuza *rain* < *qu(ŋ)ZaN

PCP *j

baja *close together* < *banzar *row*
 duji *point* < *tunzuk
 guju- *mouth* < *ŋusu *upper lip* (or *ijun *nose*)
 laja *tame* < *Najam
 majaga- (*road+*) *fork* < *saŋa
 taji *shave* ?< *ta(zZ)im *sharpen*
 uja *transport, carry (cargo)* < *(Rl)ujan (cf. Proto-Philippines *lújan *ride, load* (Zorc 1985))

The evidence now accumulated requires a revision of my tentative proposal of 1983, to include PAN *z and *Z with *j, and to include in their reflexes PCP *j as well as *z. The hypothesis now proposed is, therefore, that PAN *s and *c became PCP *c/s, and that PAN *j, *z, and *Z became PCP *z/j. Of the 55 examples above, only eight are in any way contradictory.

If the above hypothesis holds, and if Blust's (1978) claim that POC distinguished *j from the other palatal obstruents is true, then POC must have distinguished three palatal obstruents, the only mergers being of *c and *s, and *z and *Z. Moreover, given the high correlation between PAN *s and PCP *c, and PAN *c and PCP *s, there may yet be a strong case for the retention of the PAN *c/s distinction in POC.

2.6.3 PCP *ñ

PCP *ñ is reconstructed as distinct from *n because of its reflexes in Rotuman (∅/r,g,n rather than n) and Western Fijian (y/n rather than n). The evidence for PCP *ñ is shown in Table 9.

In Western Fijian, PCP *ñ becomes n before u, y before a. In Rotuman, it becomes r initially, and ∅ medially, usually fronting or raising the following vowel; the n and g reflexes appear to be sporadic. In Eastern Fijian and Polynesian, *ñ merges with *n as n.

Table 9: Evidence for PCP *ŋ

PCP	PWF	PEF	PPN	ROT
ñamu <i>mosquito</i>	yamu	namu	namu	ramu
N-ñā IIII	-ya	-na	-na	-na, -ga ¹
dañudañu <i>fallow</i> ²	danudanu	danudanu	-	taitai
mañawa <i>spirit</i>	-	-	manawa	maeva
meña (<i>breadfruit</i>) <i>ripe</i> ³	meme	-	-	mea
moña- <i>brain</i>	moya	mona	-	-
voñu <i>turtle</i>	-vonu ⁴	vonu	fону	hoi
voñu <i>full</i>	-	-	fону	hoi
waña <i>k sea-urchin</i>	-	-	wana	vaevae

NOTES

1. -na is productive, but -ga is fossilised in: *sasiga younger same-sex sibling* (*tazi-ñā), *ma'piga grandparent, grandchild* (*makubu-ñā), *uluga top, summit* (*'ulu-ñā), *laloga inside* (*lalo-ñā).
2. Apparently distinct from POC *talū *fallow*, reflected by PPN *talū.
3. External witness: PSS *me(nñ)a *ripe*.
4. Reflected in *tuvonu k turtle, Caretta caretta*.

2.6.4 PCP *y

PCP *y is realised as PFJ *c, PPN ∅, and Rotuman r-/-∅-. The fact that initial *y becomes r in Rotuman, and that intervocalic *y affects the following vowel in exactly the same way as intervocalic *ñ, suggests that PCP *ñ and *y merged as pre-Rotuman *y. The evidence for PCP *y is presented in Table 10.

Table 10: Evidence for PCP *y

PCP	PFJ	PPN	ROT
yagi <i>wind</i>	cagi	agi (<i>wind</i>) <i>blow</i>	ragi <i>breeze, breath</i>
yago <i>k plant, Zingiber sp.</i>	cago	ago	raga
yavā <i>storm</i>	cavā	afā	-
yavo <i>fishing-line</i>	cavo	afo	-
kayu <i>wood, tree</i>	kacu	lākau	?ai
laya <i>sail</i>	la(cz)a	lā	lae
maya <i>ashamed</i>	-	mā	mae
?ayawa <i>k tree, Ficus sp.</i>	yacawa	?awawa	aeva

2.7 VELARS

Table 11: The reflexes of the PCP velars

PCP	x	k	kw	q	qw	g	gw
PFJ	k	k	kw	q	qw	g	gw
PPN	ʔ	k	k	k	k	g	g
ROT	∅/?	ʔ	ʔ	k	k	n/g	v

PCP *x has not been reconstructed before. It is distinguished from *k by the reflexes PPN *ʔ (or occasionally a ʔ/k doublet) rather than *k (as reported in Geraghty 1983:160-161), and Rotuman ∅ or ʔ rather than ʔ. It may perhaps turn out to be the result of an incomplete change, rather than an actual PCP phoneme; but it is convenient at this stage to catalogue it as *x, and its inclusion lends symmetry to the system, since the velar series now parallels the labial. The evidence for PCP *x is presented in Table 12. What little Rotuman evidence there is points to the reflex ∅ before a and ʔ before u, with the reflex before o equivocal.

Table 12: Evidence for PCP *x

PCP	PFJ	PPN	ROT
INITIAL			
xa(bp)a <i>(house) wall</i>	ka(bp)a	(ʔ)apa(ʔ)apa	-
xana-N <i>past</i>	kana-	ʔana-	-
xanusu <i>spit</i>	kānusi	ʔanusu	anusu (PN?)
xata <i>make mark, show clearly</i>	kata	ʔata <i>shadow, reflection; (tattoo) bright (REN)</i>	afa
x(a)ua <i>don't, cease</i>	kua	(kʔ)aua	ʔuʔua
xavelu, vaxelu <i>wipe anus</i>	kāveʔlu	faʔelu	-
xoda <i>eat raw (flesh)</i>	koda	ʔota	-
xoia	kolia <i>split (fire-wood)</i>	ʔolia <i>wedge</i>	ʔolo <i>chop, cut</i>
xota <i>dregs, refuse</i>	kota	ʔota	mofa ¹
xōtai <i>fruit salad</i>	kōtai	ʔōtai	-
xū-t <i>bite off</i>	kū-t	ʔū-t	-
xujim(ai), xumij(ai) <i>crave fish or seafood</i>	kusima	ʔumiti	-
xuru <i>rumble</i>	kuru	ʔuʔulu	ʔū <i>bang</i>
MEDIAL			
axa- <i>jawbone</i>	yaka- <i>mouth</i>	aʔa	-
boxoi <i>k pudding</i>	bokoi	poʔoi	poʔoi (PN?)
jexejexe <i>k fish, Arothron</i>	sekeseke	teʔeteʔe	-
maxavu <i>Magellan's clouds</i>	makavu	maʔavu	-
maxota <i>k tree, Dysoxylum</i>	mākota	maʔota	-
mexe <i>dance</i>	meke	meʔe	-
noxa <i>tie up, tether</i>	noka	noʔa	-
saxalo <i>scraper (coconut)</i>	i/sakalo <i>coconut scraper</i>	saʔalo, sākalo	-
tānoxu, tāxona <i>k bowl</i>	tākona	tānoʔa	-
vaxa-V <i>often</i>	vaka-	faʔa-	-
vaxo <i>peg, nail</i>	vako	faʔo	-
NOTE			
1. Sporadic prothesis of m before back vowels in Rotuman is not unlikely; the prothesis of r and g before initial a will be discussed below.			

PCP *kw, *qw, and *gw are reconstructed as distinct from *k, *q, and *g, respectively, because of their reflexes kw (or xw), qw, and gw in Western Fijian and the south-east Vitilevu area of Eastern Fijian (see Geraghty 1983:42-50). The labiovelars (*kw, *qw, *gw) only occur before a and, far less frequently, e. PCP *gw is the reflex of POC *mw, and is distinguished also in Rotuman as v, rather than n/g from PCP *g. PCP *qw may be a conditioned reflex of PEO *bw (in addition to PCP *b (Geraghty 1983:120-124)), but the evidence is as yet only suggestive, and is not presented here. External cognates of PCP *kw are the same as those of *k. Some examples of PCP *k, *kw, *q, *qw, and *g are given below, and the evidence for PCP *gw is presented in Table 13.

k-

*kini *pinch*: PFJ *kini, PPN *kini, ROT ?ini
 *kati *bite*: PFJ *kati, PPN *kati, ROT ?afi
 *kuli- *skin*: PFJ *kuli-, PPN *kili, ROT ?uli

-k-

*kauki *sand-crab*: PFJ *kauki, PPN *kawiki, ROT ?avi?i
 *coka *husk (coconut)*: PFJ *coka, PPN *hoka, ROT so?a
 *(cz)akule *search for lice*: PFJ *(cz)akule, PPN *sakule, ROT sa?ule

kw-

*kwai *say, tell*: PFJ *kwai, PPN *kai *tell story*, ROT ?ea (?< *kwai+a)

-kw-

*sikwa *net-needle*: PFJ *sikwa, PPN *sika, ROT si?a
 *bekwa *fruit-bat*: PFJ *bekwa, PPN *peka

q-

*qiriqiri *gravel*: PFJ *qereqere (for *qiriqiri), PPN *kilikili
 *qau *swim*: PFJ *qau, PPN *kau, ROT kau *wade*
 *qumuqumu *k crab*: PFJ *qumuqumu, ROT kumkumu

-q-

*leqileqi *k tree, Xylocarpus*: PFJ *leqileqi, PPN *lekileki, ROT lekileki (PN?)
 *waqa *canoe*: PFJ *waqa, PPN *waka, ROT vaka (PN?)
 *(y)aqo *learn*: PPN *ako, ROT rako

qw-

*qwalae *k bird, Porphyrio, swamphen*: PFJ *qwalā, PPN *kalae, ROT kalae
 *qwele *earth, dirt*: PFJ *qwele, PPN *kele, ROT kele *black, blackish*

-qw-

*nuqwa *k tree, Decaspermum fruticosum*: PFJ *n(iu)qwa, PPN *nukanuka

g-

*gi(czs)a: ROT nisa *mock, jeer* (PNCV *gigica *smile, grin*)
 *gara *scream, cry loud*: PFJ *gara, PPN *gala
 *guju- *mouth*: PFJ *gusu-, PPN *gutu, ROT nuju

-g-

- *tagi *cry*: PFJ *tagi, PPN *tagi, ROT fagi
- *taliga- *ear*: PFJ *taliga-, PPN *taliga, ROT faliga
- *togo *mangrove*: PFJ *togo, PPN *togo, ROT fogo

Table 13: Evidence for PCP *gw

PCP	PFJ	PPN	ROT
gwaʔane <i>male</i>	-gwane	-gaʔane	vavane <i>husband</i>
gwa(cz)a(cz)i <i>k fish, Parupuneus</i>	g(w)a(cz)a(cz)i	-	vasasi
gwalu <i>wave, surf</i>	-	galu	valu
gwajala <i>k fish, Epinephelus</i>	kasala (*gwasala)	gatala	vajala
gwata <i>snake</i>	gwata	gata	-
dagwa <i>loose, slack</i>	dagwa	tagataga	-
regwa <i>turmeric</i>	reregwa	rega	-

2.8 GLOTTAL AND ZERO

Table 14: The reflexes of the PCP glottal and zero

PCP	ʔ	∅/#_a
PFJ	∅/y(/#_a)	y
PPN	ʔ	∅
ROT	∅/?	r, g

The above interpretation is innovative with respect to Rotuman. Biggs (1965: 408-409) claimed Rotuman simply lost PCP *ʔ, and made no reference to prothesis before *a.

All instances of PCP initial *a reflected in Rotuman show a prothetic r or, in two cases, g:

- *agi *give instructions, urge on* > ragi
- *aka *k vine, Pueraria lobata* > gaʔa
- *atu *large number* > rafu (POC *Ratu *hundred*)
- *aqo *learn* > rako
- *au *Il* > gou/a
- *ava *handle* > hara (met.)
- *avi *fire* > rahi

Forms showing initial a in Rotuman derive from PCP *ʔa or *xa:

- *ʔayawa *k tree, Ficus* > aeva
- *ʔate *liver* > afe
- *ʔatu *line, row* > afu
- *ʔatule *k fish, Selar crumenophthalmus* > afule
- *ʔa(cz)o *sun* > asa
- *ʔanuve *caterpillar* > aniha

*xanusī *spit* > anusī
 *xata *make mark, show clearly* > afa

The fate of PCP *ʔ in other environments is less clear. It is usually lost:

INITIAL

*ʔo(cz)o *provisions for journey* > oso
 *ʔoti *finished* > ofi
 *ʔulu-ña *its top, summit* > uluga
 *ʔunavi *scale (fish)* > unehi
 *ʔuta *inland* > ufa
 *ʔuza *rain* > usa

MEDIAL

*liʔo *voice* > lio
 *mataʔu *right-hand* > mafau
 *matuʔa *old* > mafua
 *raʔa *branch, twig* > rā
 *taʔo *cook* > fao
 *taʔu *year* > fau
 *tinaʔe *intestines* > finae
 *tuʔa *ridge, (leaf) midrib* > fua
 *tuʔu *stand* > fū
 *vaʔa *stalk* > hā
 *vetaʔu *k tree, Calophyllum inophyllum* > hefau

In some cases, however, PCP *ʔ appears to be retained:¹¹

*ʔuvi *yam* > ʔuhi
 *ʔo- *possessive (PPN *(ʔ)o- (Wilson 1982:73))* > ʔo-
 *ʔi *at, in, on* > ʔe
 *suʔi *pour water on* > suʔi

In the light of changes posited earlier in this paper, there now appear certain parallel developments which are best explained by two important mergers in the early history of Rotuman.

(1) PCP *ʔ, *x, and some cases of *v (probably via *h) merged as pre-Rotuman *ʔ. This phoneme must have been present at the time of r/g prothesis. Subsequently, *ʔ was lost before a, but retained in some cases before other vowels, especially u.

(2) PCP *ñ, *y, and *∅/#_a merged as pre-Rotuman *y, rarely *g. Subsequently, *y became r initially. Medial *y changed following a to e and u to i, and then was lost.

NOTES

1. Hockett (1976) reconstructed a phonology and a large number of lexical items for "Proto Fiji Polynesian", the language ancestral to Fijian and Polynesian, but did not consider Rotuman.
2. Pawley has claimed (1979:13) that there is enough evidence to support a Rotuman-Fijian subgroup exclusive of Polynesian.
3. There are also problems relating to the reality of Proto-Fijian - Geraghty and Pawley (1981) have suggested that some features now widespread in both major subgroups of Fijian, Eastern and Western, may be the result of diffusion after the break-up of Proto-Fijian.
4. Failure to see this development led Biggs (1965) to posit Rotuman *f* as the regular reflex of PCP **nt* in initial position. The examples cited were all in fact of PCP **t*.
5. A five-vowel system, with phonemic length, is also indicated by the evidence. Its development, though not entirely straightforward, will not be discussed in detail here. A major feature of Rotuman is the lowering of final **o* and **e* to *a* (noted in Pawley 1979), under conditions yet to be determined.
6. Vowel lengthening is a common sporadic development in Proto-Fijian. Parallel to this example are **mīmī* *urinate* (< **mimi*) and **qōqō* *narrow* (< **qoqo*). PFJ **vāō* *k tree*, *Ochrosia* (< PEO **vaRo*) seems to result from avoidance of the vowel cluster **ao* (Geraghty and Pawley 1981), instead of the usual simplification. Especially common is the lengthening of pretonic **a*, as in PFJ **kānusi* *spit*, **kāvelu* *wipe anus*, **q(w)ālotu* *egg*, and **mā(cz)awa* *space between*.
7. Other examples of final long vowels shortened in PPN and Rotuman:
 **dulī* *plover*: PPN **tuli* (but SAM *tulī*), ROT *juli*
 **takū*- *back*: PPN **tak(uū)* *tortoiseshell*, ROT *faʔu*
 **vetuʔu* *star*: ROT *hefu*
 **tulū* *k small land crab*: PPN **tulu*
8. A similar assimilation has applied in *nunami k edible sea-anemone* (< **drumani*, with metathesis), and *hahiʔa Malay apple* (< **kavika*); and optionally in the loanword *saujia, jaujia soldier*.
9. PCP **v* is sporadically reflected as **b* in PFJ (or parts of Fiji), e.g.: *bā taro stem* (< **vaʔa*), *bō squeeze* (< **vō*), *buka firewood* (< **vuka*).
10. Probably also Levy's PSS **j*, which seems to me may turn out to be a conditioned reflex of **d*. Note also my suggestion (Geraghty 1983:193) that PCP **j* has a distinct cognate in PSS, because the Sandfly Passage dialect of Nggela, according to Fox (1955), often shows *s* for *d*, and the cases cited include cognates of PCP **j*, not of PCP **d* or **dr*.
11. These, however, may not be genuine retentions, but cases of intrusive glottal stop, as shown in these forms:
 **ifi* *k tree, Inocarpus fagiferus* > ?*ifi* (PN)
 **jamua* (*palm*) *fruit stem* > *jamuʔa*
 **mea reddish* > *mea, miʔa* (PN?)
 **tea white, pale* > *fea, feʔa*
 **tulou* *word of apology* > *turoʔu* (PN)

REFERENCES

BENDER, Byron W. et al

- n.d. Proto-Micronesian word-list. Computer printout. Department of Linguistics, University of Hawaii.

BIGGS, Bruce

- 1965 Direct and indirect inheritance in Rotuman. *Lingua* 14:383-415.
 1978 The history of Polynesian phonology. In Wurm and Carrington, eds 1978:691-716.
 1979 *Proto-Polynesian word-list II*. Working Paper No.53, Department of Anthropology, University of Auckland.

BLUST, Robert A.

- 1976 A third palatal reflex in Polynesian languages. *Journal of the Polynesian Society* 85/3:339-358.
 1978 *The Proto-Oceanic palatals*. *Journal of the Polynesian Society* Monograph No.43. Auckland.

CHURCHWARD, C.M.

- 1940 *Rotuman grammar and dictionary*. Sydney: Australasian Medical Publishing Company.

CLARK, Ross

- 1985 Languages of north and central Vanuatu: groups, chains, clusters and waves. In Andrew Pawley and Lois Carrington, eds *Austronesian linguistics at the 15th Pacific Science Congress*, 199-236. PL, C-88.

DEMPWOLFF, Otto

- 1934-38 *Vergleichende Lautlehre des austronesischen Wortschatzes*. *Zeitschrift für Eingeborenen-Sprachen*, Beiheft 15 (1934), Beiheft 17 (1937), Beiheft 19 (1938). Berlin: Dietrich Reimer.

ELBERT, S.H.

- 1953 The internal relationships of Polynesian languages and dialects. *Southwestern Journal of Anthropology* 9/2:147-173.

FOX, C.E.

- 1955 *A dictionary of the Nggela language*. Auckland: Unity Press.

GERAGHTY, Paul

- 1983 *The history of the Fijian languages*. Oceanic Linguistics Special Publication No.19. Honolulu: University of Hawaii Press.

GERAGHTY, Paul and Andrew K. PAWLEY

- 1981 The relative chronology of some innovations in the Fijian languages. In Hollyman and Pawley, eds 1981:159-178.

GRACE, George W.

- 1959 *The position of the Polynesian languages within the Austronesian (Malayo-Polynesian) language family*. Memoir 16, International Journal of American Linguistics. Bishop Museum Special Publication 46.

- 1967 Effect of heterogeneity in the lexico-statistical test list: the case for Rotuman. In Highland et al, eds 1967:289-302.
- GREEN, Roger and M. KELLY, eds
- 1972 *Studies in Oceanic culture history*, vol.3. Honolulu: Bishop Museum.
- GUY, Jacques Bernard Michel
- 1978 Proto-North New Hebridean reconstructions. In Wurm and Carrington, eds 1978:781-850.
- HALE, Horatio
- 1846 *Ethnography and philology. United States exploring expedition during the years 1838, 1839, 1840, 1841, 1842 under the command of Charles Wilkes*. Philadelphia: Lee and Blanchard.
- HAUDRICOURT, André-Georges and Françoise OZANNE-RIVIERRE
- 1982 *Dictionnaire thématique des langues de la région de Hienghène*. Paris: Laboratoire des Langues et Civilisations à Tradition Orale.
- HIGHLAND, G.A. et al, eds
- 1967 *Polynesian culture history: essays in honor of Kenneth P. Emory*. Bishop Museum Special Publication 56. Honolulu.
- HOCKETT, Charles
- 1976 The reconstruction of Proto Central Pacific. *Anthropological Linguistics* 18/5:187-235.
- HOLLYMAN, Jim and Andrew K. PAWLEY, eds
- 1981 *Studies in Pacific languages and cultures in honour of Bruce Biggs*. Auckland: Linguistic Society of New Zealand.
- LEVY, Richard
- n.d. Languages of the Southeast Solomon Islands and the reconstruction of Proto Eastern Oceanic. Department of Anthropology, University of Kentucky.
- MILKE, Wilhelm
- 1961 Beiträge zur ozeanischen linguistik. *Zeitschrift für Ethnologie* 86:162-186.
- 1968 Proto-Oceanic addenda. *Oceanic Linguistics* 7:147-171.
- PAWLEY, Andrew K.
- 1972 On the internal relationships of Eastern Oceanic languages. In Green and Kelly, eds 1972:1-142.
- 1979 *New evidence on the position of Rotuman*. Working paper No.52. Department of Anthropology, University of Auckland.
- RANBY, Peter
- 1980 *A Nanumea lexicon*. PL, C-65.
- TURNER, George
- 1884 *Samoa*. London: Macmillan.

WALSH, David S.

- 1984 A presentation of some Proto-North-East Vanuatu reconstructions, and a consideration of their historical implications. Paper presented at the Fourth International Conference on Austronesian Linguistics, Suva, Fiji.

WILSON, William H.

- 1982 *Proto-Polynesian possessive marking*. PL, B-85.

WURM, S.A. and Lois CARRINGTON, eds

- 1978 *Second International Conference on Austronesian Linguistics: proceedings*. PL, C-61.

ZORC, David

- 1986 The genetic relationships of Philippine languages. In this volume, 147-173.

THE CHRONOLOGY OF THREE SAMOAN SOUND CHANGES

Even Hovdhaugen

1. INTRODUCTION¹

In Buck 1930:4-5 we find the following observations on Samoan historical phonology:

Recent changes have taken place in the spoken language in the substitution of *k* for *t* and a loose mutual interchange between the sounds *n* and *ng*. The reintroduction of *k* in place of *t* is extremely interesting as it evidently indicates a Polynesian tendency not confined to one dialect. A similar change has already completely occurred in the Hawaiian dialect in which it passed through two distinct phases. Thus, in the widespread Polynesian word *kumete* (wooden bowl) the first phase was the dropping of the *k* so that the word became *'umete*. In the second phase which occurred later, the *t* was changed to *k* and the word became *'umeke*. Thus the lost *k* came back into the dialect but in no word did it re-occupy its original position. In the process of resurrection, the *k* displaced the *t* sound completely out of the dialect. In Samoa, the first phase of dropping the *k* had been completed before the Bible was printed in Samoan and *kumete* had become *'umete*. The second phase of substituting the *k* for *t* is now taking place in everyday speech and a wooden bowl is now more referred to as *'umeke* than as *'umete*. The talking chiefs make the change in official speeches and the retention of the *t* sound is regarded by the public as pedantic. It seems probable that the Samoan *t* like the Hawaiian *t* is doomed to extinction.

The interchange between *n* and *ng*² has become so common that I had to constantly consult Pratt to find which was the original sound used. Thus in spoken speech, it is more usual to hear *paono* instead of the correct *paongo*, and *tafangi* instead of the correct *tafari*.

Buck's treatment does not differ substantially from that of other scholars who have touched on these problems, except maybe that Buck's presentation is a bit more shaded and above all more precise concerning the chronology of the sound changes than other treatments of them are.

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 313-331. *Pacific Linguistics*, C-94, 1986.

© Even Hovdhaugen

2. THE DEVELOPMENT PP³ *k > ' ⁴ IN SAMOAN

There are strong reasons for assuming that this sound change took place before the first contacts between Samoans and Europeans. W. Anderson, the surgeon on *Resolution* on Cook's third voyage to the Pacific mentioned in his journal of July 1777 three Samoan words which he had picked up on Tonga and "which seem to differ much from those of the other islands" (Beaglehole 1967:958). One of these words is tamaë'ty a *chief woman* which is identical with Mod. Sam.⁵ tama'ita'i *lady*.

Since the etymology of tama'ita'i isn't clear, the evidence provided by this example may not appear conclusive, but that objection does not apply to the material provided by Captain Edwards and surgeon Hamilton on *Pandora* when they visited Samoa 14 years later (1791). There we find forms like Oattooah (Thomson 1915:50) = Mod. Sam. 'O Atua and Ootootoillah/Otutuelah (Thomson 1915:55,130) = Mod. Sam. 'O Tutuila both containing the Samoan topic marker 'o < PP *ko.

Especially important when evaluating the information provided by Anderson, Edwards and Hamilton is the fact that on both expeditions there were several people familiar with Polynesian languages and accordingly capable of obtaining trustworthy information from the native informants. That was not the case with many other visitors and the materials provided by them are of very variable quality.

First of all we have the reports from La Pérouse's tragic visit to Samoa in 1789. The main source for this expedition is Milet-Mureau 1797, III:227-238 and there we find the following Samoan place names mentioned: Opoun, Léoné, Fanfoué, Maouna, Oyolava, Pola, Calinassé, Shika, Ossamu and Overa. Nobody on La Pérouse's expedition was familiar with Polynesian languages and the general information about the Samoan language in the report as well as the qualifications of the "interpreter" do not support the reliability of the examples mentioned:

Nous n'avions d'abord reconnu aucune identité entre leur langage et celui des peuples des îles de la Société et des Amis, dont nous avons les vocabulaires; mais un plus mur examen nous apprit qu'ils parlaient un dialecte de la même langue. Un fait qui peut conduire à le prouver, et qui confirme l'opinion des Anglais sur l'origine de ces peuples, c'est qu'un jeune domestique manillois, né dans la province de Tagayan au Nord de Manille, entendait et nous expliquait la plus grande partie des mots des insulaires: on sait que le tagayan, le talgale, et généralement toutes les langues des Philippines, dérivent du malais et cette langue, plus répandue que ne le furent celles des Grecs et des Romains, est commune aux peuplades nombreuses qui habitent les îles de la mer du Sud.

(Milet-Mureau 1797, III:229-230)

The Samoan sources for the names given by La Pérouse are still a matter of pure guessing (Oyolava = 'O i ō lava *Just there* ?), cf. Krämer 1903:13 and Von Bülow 1900 for further discussion.

The next visitor to provide some linguistic material of interest is Otto von Kotzebue (1830). But his material is no less confusing than that of La Pérouse and the linguistic competence regarding Polynesian languages was also practically nonexistent on Kotzebue's expedition. He mentions that the Samoan word for *pig* is boaka (Kotzebue 1830:279), cf. Mod. Sam. pua'a, PP *puaka, but as Krämer (1903:19) had already pointed out, Kotzebue probably had contact with some Tongan

speaking people in Samoa (cf. that the Tongan word for *pig* is *puaka*) since he refers to the name for a *chief* as *eige* which clearly is Tongan 'eiki and not Samoan ali'i. Let it also be added that Kotzebue generally has been considered by posterity to be a rather superficial and bad observer, cf. Krämer 1903:17 and the more or less justified attacks on him by the LMS missionaries (Williams 1837: 482 and especially Ellis 1831).

In Williams 1837:451 we find an interesting form pertaining to the problem we are investigating. A chief of Malova is reported to have made the following statement:

I am now a worshipper of Jehovah, my heart and thoughts are in love with the good word, and my sincere desire is that, speedily, it may spread through the land, and that not a *Faka-Devololo*, a devil's man, may remain.

But the expression *Faka-Devololo* (Mod. Sam. fa'atevololo) may be another example of Tongan influence, having come into use in Samoa through the teaching of the Wesleyan Tongan missionaries there, through Samoan visitors to Tonga having picked up the expression there or through Tongan visitors to Samoa. As we will see below there are other indications of such Tongan influence on Samoan.

The last example I have found of PP *k apparently being retained as k in Samoan is in Walpole 1849, II:331 where he mentions that when approaching Upolu:

Our vessel was soon surrounded by canoes, laden with their various treasures, the natives shouting "Fackaton, fackaton" (barter, barter)

The word *fackaton* most likely represents Tongan *fakatau*, cf. Mod. Sam. fa'atau.

The only conclusion we can draw from these scattered examples is that they are of little if any importance for the chronology of the change *k > ' since they could all be Tongan loanwords. But they are interesting in so far as they indicate that a certain familiarity with the Tongan language existed in Samoa in the early 19th century and that the use of Tongan phrases and words was to some extent common and perhaps fashionable.

The connections with Tonga lead us to another interesting set of data, namely the two loanwords 'oti *goat* (< Eng., cf. Tong. kosi < *koti⁶) and *tapa'a tobacco* (< Eng., cf. Tong. tapaka). These words cannot be older in Samoan than the knowledge of the objects they designate and as far as goats are concerned they were first known in this part of the Pacific when Captain Cook introduced goats to Tonga on his third voyage (cf. Beaglehole 1967:134,155), on exactly the same voyage when Anderson noted examples indicating that the sound change *k > ' already had taken place in Samoan.

The explanation must be sought in the fact that there were at that time extensive contacts between Samoans and Tongans. The examples referred to earlier in this paper indicate a certain knowledge of Tongan in Samoa and the use of Tongan in idioms, slang or other contexts. Further, the earliest reports from the LMS and especially Wesleyan missionaries in Samoa are full of references to Samoans having visited Tonga, Tongans staying in Samoa and the ease with which the Wesleyan missionaries from Tonga conversed with Samoans compared to the problems their Tahitian and Rarotongan colleagues in the LMS had in being understood.

Another important source is William Mariner's account of his stay in Tonga 1806-1810. He has several references to the high status of Samoan in Tonga and the popularity of Samoan songs there (Martin 1817, I:165-166,375, II:228-229,323,

337-339). The specimens he quotes of such songs are unfortunately of little value since the texts are too distorted to be intelligible for the simple reason that Mariner did not understand Samoan as he frequently pointed out and "very few of the Tonga people understand the Hamoa language" (Martin 1817, II:323). Nevertheless, Mariner tells us explicitly that he has got his information about Samoa from Samoans in Tonga (e.g. Martin 1817, I:1375) and that must imply that those Samoans spoke Tongan to him.

The linguistic contact between Tongans and Samoans must have made both parties aware of some of the basic and quite regular sound correspondences between the two languages. Accordingly, knowing that Tongan *k* regularly corresponded to Samoan ' it is likely that Samoans substituted a ' for a *k* in borrowings. But it is also likely that this substitution implies that Samoan at that time did not have any phoneme /*k*/ and not even the sound [k].

There are also a few other interesting words which may be loanwords of the same category and period as 'oti and tapa'a:

u'amea *iron, metal*, cf. Tongan ukamea/'ukamea⁸ and Fiji kaukamea.

'ulo *pot*, cf. Tongan kulo⁹ and Fijian kuro.

None of these words have Polynesian cognates outside the Tongan-Fijian-Samoan triangle (Niue *ulo pot* is a late borrowing from Samoan introduced by the missionaries) and already Pratt 1862:95 (cf. also Pratt 1977:67) considered 'ulo to be a recently introduced word in Samoan. Neither iron/metal nor pots were elements of traditional Samoan culture and we do not expect to find indigenous terms for them.

It is also interesting to observe that the word for cooking in pots is saka in Samoan, a word generally acknowledged to be borrowed from Fijian, cf. Fijian saqa.¹⁰ And this is one of the very few Samoan words where *k* is obligatory even in Lit. Sam., cf. below.

While 1777 then seems to be the terminus ante quem for the development **k* > ' , the question concerning a possible terminus post quem is much more difficult and for the moment probably impossible to answer. We do not have any written sources before 1777 which give us any data concerning Samoan and the use of oral sources and traditional material for diachronic linguistics is difficult and has obvious limitations. Let me take but one example which as far as I know is also the most relevant one for the problem we are discussing. If the Rarotongan hero Karika came from Samoa (cf. Gill 1876:25), if he is identical with the chief 'Ali'a of Manu'a, if we trust the Rarotongan genealogies (cf. Browne 1897:10) and assume that he lived in the 15th century, then it is reasonable to assume that Samoan still had the original *k* retained at that time since 'Ali'a most probably would have been rendered as *'Ari'a in Rarotongan. But there are a few ifs too many in this argumentation to base any conclusions on it.

As a last resort we may try to look for comparative evidence for the chronology, but the results are meagre. Among Western Polynesian languages only Samoan and Luangiua have the sound change **k* > ' and in these two languages the changes are completely parallel, cf. below concerning other cases of parallelism between the two languages. Similar sound changes in Eastern Polynesian seem to have no connection with the Samoan one: in Tahitian we have a general changing of velars (i.e. *g* and *k*) into ' and in Hawaiian the change **k* > ' is still not completed and there are several lexemes where the *k* is retained as an optional variant, cf. Elbert 1982:503-504. The intricate dialectal division and intermingling in Marquesan (cf. Elbert 1982), which by Polynesian standards is unique, has only been superficially described, but according to the information available it seems

that *k > ' is restricted to some dialects, *g > ' is only sporadic and restricted to some lexemes, while one of the most characteristic features of Marquesan is *r and *l > ', a development not attested elsewhere in Polynesia.

3. THE VELARISATION OF /n/ AND /t/ IN SAMOAN

From a phonological point of view we find two distinct sociolects in Samoan today¹¹ characterised by the following two different sets of nasal and plosive phonemes:

Literary Samoan	Colloquial Samoan
p t k	p k
m n g	m g

One can formalise the correspondences between the two systems in the following rather simple rule, assuming that the system of Lit. Sam. is the older one:

$$+ \text{ dental } \text{---->} + \text{ velar } / \left[\begin{array}{l} \propto \text{ plosive} \\ - \propto \text{ nasal} \end{array} \right]$$

But the development /t/ > /k/ and the development /n/ > /g/ are on neither the synchronic nor the diachronic level identical. Both /n/ and /g/ are full-fledged phonemes in Lit. Sam. representing generally PP *n and *g while /k/ in Lit. Sam. only occurs in recent borrowings from English, e.g. keke *cake*, kofe *coffee*, māketi *market* and in a few other words like puke,¹² 'expression to startle someone', puketā, 'exclamation of triumph in the game tāgāti'a', pukē *basket (used for storing cloth)*, saka *cook*. A few words which formerly had t in Lit. Sam. are nowadays considered by most speakers to have a basic k, e.g. okaoka (otaota), 'exclamation of surprise, shock', oka (ota) *raw fish*.

Accordingly, the development t > k is to a great extent a change in the phonetic realisation of one and the same phoneme and only in a subfield of the lexicon does it influence the phonological system and the set of distinctive oppositions. Samoans have generally no problems in switching between the two sociolects as far as plosives are concerned and errors like t-forms in k-style and vice versa are very seldom attested.¹³ It is quite to the contrary with the development n > g. In that case one has to know by heart which words have which phoneme(s) and most Samoans have great problems in separating the two styles with regard to nasal phonemes. The number of hypercorrect n-forms (instead of basic g-forms) in attempts at speaking Lit. Sam. is significant and so is the number of colloquial g-forms in literary style. This mixture of the two styles is not restricted to the spoken variant but is also frequently encountered in written sources.¹⁴

This intermingling of the two styles as well as the strong degree of variability in Spoken Lit. Sam. is hardly mentioned in the existing grammars, probably due to the fact that their main source of data is the classical literary language found in the Bible and similar sources. Editions of Samoan texts are also in general strictly normalised (this even applies to publications of typical oral texts like Stuebel 1895 and Moyle 1981). One of the few exceptions to this is Sierich 1901-1903 whose texts give a faithful depiction of the mixture of the two styles.

There are various theories about the origin of the sound changes t > k and n > g, but all of these seem to consider them as one simultaneous sound change

and not as two separate processes. Further they are considered as a quite recent phenomenon (or corruption as most would say) taking place in the last part of the 19th century when the literary Samoan language was well established.

In Pratt 1862 these changes are not mentioned at all and it is said explicitly that *k* is an introduced foreign letter used for writing proper names and loanwords (Pratt 1862:6).¹⁵ Among examples of careless pronunciation of consonants *lagoga* for *lagona* is mentioned, but it is not indicated that it is a general feature of the language. In a later edition of the grammar he has a note in the preface dated 5 June 1876¹⁶ saying:

In Hawaii they have changed the *t* into *k*, and *ng* into *n*. Thus *tangata* has become *kanaka*. Samoans are doing the same thing at the present time, to the great injury of the language.

An interesting treatment of these changes which sums up most of the current views at the turn of the century is Von Bülow 1897a:375:

Bezüglich der Änderung - Verderbnis nenne ich es - des *t* in *k*, des *n* und des *g* in *ng*, sagt ein Kenner der Samoasprache, der Missionar S.J. Whitmee, in einer Anmerkung zu Pratts Wörterbuch der Samoasprache (S.1): "This is a recent change. When I went to Samoa in 1863 I heard *k* used only on the island of Tutuila and on the eastern portion of Upolu. Now it is used all over the group. It is difficult to say how this change commenced, but its spread has been noted. - - The more intelligent use *t* quite correctly in reading and in public speaking. But the practice of transposing *k* and *t* in reading is rapidly growing." -

Selbst in Fremdwörtern, in denen *k* vorkommt, vertauschen es die Eingeborenen jetzt sehr oft mit *t*, und umgekehrt. In derselben Weise werden *n* und *g* (*ng*) miteinander vertauscht.

"Both of these changes took place in the Hawaiian dialect at a much earlier date and they have been adopted in the literature (sic!) of the Hawaiian Islands, which is not the case in Samoa".

Krämer 1897:77 has the following more laconic view of the situation: "Die Aussprache des *k* für *t* ist von Hawaii eingeschlept. Auf Savaii wird indessen meist noch das reine *t* gesprochen".

Brown 1916:182 gives a rather detailed description of the change *t* > *k* (*n* > *g* not being mentioned) which in some respects deviates from other sources:

When I first resided in the group, in 1860, there were very few people indeed on Upolu, and still fewer, if any at all, on Savaii who used the "*k*" sound. It was very rarely heard outside of the Port of Apia and the Tuamasaga district.

The general opinion was that the change was introduced from the island of Tutuila, and it was certainly called *O le nanu faa-Tutuila* (Tutuila jabber, or wrong speech). Whether it originated on Tutuila or not I cannot say. Some individuals seemed to be conscious of their wrong pronunciation. One man tried to excuse himself by saying that his mouth was hard. The spread of the change was very gradual,

and I am inclined to believe that intercourse with white men hastened the progress of it, for many of the traders used the pronunciation.

There are some points connected with this innovation which may be noticed:

1. According to the testimony of the natives the sound of "k" was not heard in the Samoan languages, except, it may be, in a very few individual cases, prior to the years 1858-60.
2. It was said to have originated in Tutuila, but none of the individuals whom I knew in the early sixties, who used the objectionable pronunciation, had ever been to Tutuila, from where the *nanu* (jabber) was said to have come. The worst case which I knew was that of a native of Manono. I do not remember hearing it on Savaii, except in a very few instances, or from visitors, whilst I was resident on that island. It certainly was not frequently used.

But Brown may not be a completely trustworthy source since he showed a very strong antipathy towards the sound change in question which he called "a very regrettable change" (Brown 1916:181) and "a threatened deterioration" (1916:183).

A modern and slightly modified version of these views is Buse 1961:105-106:

The following is perhaps the origin of the two styles. In the nineteenth century, there began in Samoan an isolative dental-velar consonant shift by which [t],[n] moved back to [k],[ŋ]. Either before the shift started or before it became at all widespread, missionaries and native Tahitian teachers arrived from Tahiti. They reduced the language to writing and set up schools, using [t] and [n] both as a basis for the script and in giving oral instruction. (Note that [t] and [n] both occur in Tahitian, but not [k] and [ŋ].) The conservative forces of writing and education were thus thrown behind the dental style and succeeded in preserving it in those fields where their influence was strongest (schools, churches, etc.). Elsewhere, however, the shift was carried through and the velar style became widespread as a colloquial medium. Further situational differentiation has set in, each style having its own fairly well-defined sphere of influence, but the position is still far from stable.¹⁷

To sum up, the scholars mentioned so far seem to assume that the sound changes $t > k$ and $n > g$ were simultaneous¹⁸ and that they:

- (a) started in the last half of the 19th century¹⁹
- (b) started in the eastern islands (Tutuila and Upolu, Manu'a not being mentioned) and spread westwards. The changes had not yet reached the western parts of Savai'i at the end of the 19th century²⁰
- (c) were originally due to influence from the Hawaiian language upon Samoan, although nobody has explained how this influence could have taken place and how the not negligible geographical distance between Hawaii and Samoa could have made contact between the two languages possible.²¹ Brown 1916:182 had the rather strange idea that white men, especially traders, promoted the change $t > k$, cf. above.

4. A CRITICAL REVALUATION OF CURRENT HYPOTHESES

Recently Shore 1982:269 has given a rather different or at least quite modified version of the story, without, however, giving very much data to support his conclusions:

Samoans sometimes claim that formal pronunciation style is still common in those areas where Samoan culture is least affected by European contact, an observation that appears almost ironic in light of the actual distribution of social contexts in which both styles are used.²² The island of Savai'i, particularly the remote village of Falealupo at the northwestern tip of the island, is held to be a bastion of "proper" (i.e. formal-style) pronunciation. My own visits there, however, did not confirm this assertion. Milner (1966) claims to have found the formal pronunciation generally used only by some older residents of Fitiuta village on Manu'a.²³

.... My best guess is that the dual phonological system in Samoan was the product of a phonological drift that had already begun by the time the first missionaries arrived in 1830. When they orthographized the Samoan language as the initial step in translating the Bible, the missionaries selected the older and culturally preferred forms, thereby "freezing" the [t] and [n] into the written language.

Although Shore's view is closer to mine than for example the theories put forward by Von Bülow or Krämer, I think it needs to be modified, and above all we have to state clearly what we really know and separate that from what is likely, sensible or possible. Let us start with the chronology.

There are few examples of k/g-style before 1850, but enough to prove that it existed. In Lundie 1846:79 we find a letter of 16 March 1840 where he tells us that when arriving in Pagopago the boat deck was covered by Samoans greeting the "Mishingalies" (= misionare in Mod. Sam.). In the first Samoan book printed (Williams 1834:5)²⁴ we find alternating forms like Fafine lalagna/Fafigne faifaiva (= Mod. Sam. fafine lalaga/fafine fai faiva). But especially important are the observations found in the journals of the LMS missionaries. Most important is the following note at the end of Buzacott's journal of 1836/37:

We have been surprized to find a number of Samoans, who cannot distinguish any difference of sound in the k & t & ng & n & who in the language use nothing but the k & ng where the t & n ought to be. It is surprizing what a difference the change of these two letters make in the beauty of the language - from being the most musical & pretty it becomes as rough & dissonant as any dialect of the South Seas.

John Williams has in his published works, as well as in the journals from his important visit to Samoa in 1832 or from later visits there, made no reference whatsoever to the existence of a k/g-style, although observations on language are far from absent in his works. Writing in Samoan or using quotations from Samoan he is also (with the exceptions noted above in Williams 1834) quite consistent in using the t/n-style. But in the journal from Williams' and Barff's first visit to Samoa in 1830 we find a very interesting observation on the language:

The language of the Samoans is a mixture of three different dialects. The Tahitian, which it strongly resembles in many words. The Rarotongan which it resembles in the nasal consonants gn etc., and the Tongatabooan which it resembles in the use of the C and K. In addition to the above they make great use of the S, which is not used by any of the above mentioned Islands.

A sensible interpretation of this is that Williams perceived velar plosives as something characteristic when first listening to spoken Samoan.

Accordingly there seems to be no doubt that the differences between the two styles were fully established about 1830. If the conclusion drawn above that Samoan had no k at the time when the words *tapa'a* and *'oti* entered the language can be accepted, then the changes $t > k$ and $n > g$ must have taken place in the period 1777-1830.

As the material quoted above shows, the current view until quite recently was that the changes started in Upolu and Tutuila and that only in this century did they reach the western part of Savai'i. Milner 1966 and Shore 1982 (cf. above) have pointed out that this view is in conflict with the present situation where the k/g-style is less common in Manu'a, the most easterly part of the Samoan islands, than elsewhere. My own field studies point in the same direction. The only Samoans I have heard using t/n-style in an informal situation when talking to other Samoans are people from Manu'a. On the other hand, the k/g-style is nowhere as prominent as in the western parts of Savai'i and many people there are hardly able to talk in the t/n-style, managing at best to produce a mixture of the two styles. The present situation is hard to conciliate with the information from late 19th century sources, referred to above, and it is hardly possible to go further than to admit that we know nothing for certain about the geographical place of origin of these sound changes and how and in what direction they spread.

But I think we can say quite a lot more about why the t/n-style was chosen as the basis of the literary Samoan language and why we in the middle of the 19th century have next to no information concerning the existence of a k/g-style in Samoa.

When the LMS missionaries after many years of struggle and hours of desperation finally had broken the code and managed to learn Tahitian and to establish a literary Tahitian language, they hoped in the beginning to get as much as possible out of this new achievement and to extend the use of it to other islands, cf. Williams 1837:122. When John Williams and Charles Barff came to Samoa in 1830, they talked to people in Tahitian or Rarotongan, they taught them Tahitian hymns²⁵ and they left behind Tahitian and Rarotongan teachers and books. The first years of the LMS' mission in Samoa was based on the Tahitian language and many of the Tahitian teachers left behind learned no Samoan or at best a rather broken Samoan.²⁶ Accordingly it would be very strange if the missionaries, when they had a choice, did not pick out the variant of Samoan phonetically closest to Tahitian as the basis for the written language and their preaching.²⁷ Whatever the social status of that variant was before the arrival of the missionaries it necessarily had to receive a high status from its position as the language of the church. Most probably both the number of speakers using it, the range of contexts it was used in and its geographical distribution increased with the spread of Christianity. It is impossible to prove, but very likely that the use of the k/g-style was more common in 1830 than in 1850 for the reasons just indicated and that explains the otherwise surprising overlooking of the colloquial

style in sources from the period 1840-1870. Maybe a more negative attitude towards Europeans and the turbulent state of affairs in Samoa at the end of the century promoted a new rise of the k/g-style. Sound changes have sometimes quite a lot in common with changes of fashion.

As Milner 1966 and Shore 1982 have pointed out the Samoans today are rather negative towards the k/g-style and to use a fashionable sociolinguistic term, the use of k/g-forms is in many contexts strongly stigmatised in Samoan. This may indicate that the difference also originally was determined by the status of person spoken to or the context, as it is today. But I think that the basis for such a conclusion is rather fragile. As we have seen above, both Pratt, Whitmee, Von Bülow and others were very negative towards the use of the k/g-style and other European authors from the 19th century are no less negative - so Funk 1893:1 says that instead of the "t" "sprechen die meisten Samoaner leider das k der Tonganer, was den gefälligen Klang der Sprache sehr beeinträchtigt". Even more outspoken is Neffgen 1904:II when he says that the pronunciation g for n is "fehlerhaft und hässlich" and that k instead of t makes "einen hässlichen Eindruck". If such attitudes were current among Europeans in Samoa and above all among the missionaries,²⁸ it is rather likely that they were adopted by the Samoans. This negative attitude towards velars was nothing peculiar to Samoa, it was something which was part of the European attitude towards all the languages of the South Seas. The works of the early explorers are full of such evaluations, but let me just quote one, which is of special relevance to the problems we are dealing with here, namely the following extract from Banks *Endeavour Journal* (Beaglehole 1963:372) from Cook's first visit to the Pacific:

All the Isles I was upon agreed perfectly as far as I could understand them; the people of Ulietea only chang'd the t of the Otahiteans to a k, calling Tata which signifies a man or woman Kaka, a circumstance which made their Language much less soft. The people of Ohiteroa as far as I could understand their words which were only shouted out to us seem'd to do the same thing, and add many more consonants and harshness's which made their Language still more untunable.

To understand this attitude properly it must be added that the Tahitian language was set up as a model of a beautiful, melodious and soft language by the first explorers and any deviation from it was considered as a deviation from perfect beauty and harmony.

5. THE SAMOAN SOUND CHANGES IN A POLYNESIAN PERSPECTIVE

The sound change PP *t > k is attested in several Polynesian languages: Luangiua, Hawaiian, Coll. Sam., some dialects of Tahitian (cf. Biggs 1978:714) and seems according to the information provided by Banks (cf. above) to have taken place on Raiatea before the native dialect there was replaced by Tahitian. Biggs (1978:712) assumes that the *t moved back to occupy the velar position left vacant by the *k > ' change. The development n > g in Coll. Sam. and Luangiua is then seen as a consequence of the first change, the velarisation of plosives being extended to nasals. This theory does not, however, explain the development g > n in Hawaiian and the rise of the phonological opposition /t/-/k/ in Mod. Sam.

Let us as our starting point take the Proto-Polynesian system of plosive and nasal phonemes:

- (1) /p t k/
/m n g/

After the *k > ' change we get the following asymmetrical system:²⁹

- (2) /p t/
/m n g/

In this system the opposition dental-velar is no longer relevant in the plosives, but still is in the nasals. Accordingly the /t/ phoneme could now be realised optionally (= stylistic variation) as [t] or [k] and we get the following situation:

- (3) Phonemes /p t m n g/
 | | \ | | |
Phonetic realisation [p t k m n g]

Depending on the frequency or the stylistic function of the non-labial plosive allophones the phoneme they represent could be conceived as basically dental (= (3)) or velar (4):

- (4) Phonemes /p k m n g/
 | / | | |
Phonetic realisation [p t k m n g]

In (3) a further development would be to extend the rule that a plosive dental could be realised optionally as either dental or velar to the dental nasal phoneme, while in (4) such an extension would affect the velar nasal phoneme:

- (3a) Phonemes /p t m n g/
 | | \ | \ |
Phonetic realisation [p t k m n g]

- (4a) Phonemes /p k m n g/
 | / | | / |
Phonetic realisation [p t k m n g]

Phonological systems like (3a) and (4a) are vulnerable since the lack of one-to-one correspondence between basic segments and phonetic realisations (= between input and output) easily leads to reanalysis and simplification. One way to go was to reestablish the one-to-one correspondence by eliminating the ambiguous phonemes:

- (3b) Phonemes /p t m g/
 | | \ | |
Phonetic realisation [p t k m g]

- (4b) Phonemes /p k m n /
 | / | | | /
Phonetic realisation [p t k m n]

(3b) is not attested in any Polynesian language of today but I think it not unlikely that this system existed earlier in Luangiua. (4b) is the system of

Hawaiian. According to the information available the Samoan system around 1830 was (3a), cf. above. But then the phoneme /k/ was reintroduced into Samoan through loanwords from Tongan, Fijian and English (perhaps facilitated through the widespread knowledge of Tongan in Samoa at that time). This reintroduction was then possible because Samoan already had the sound [k] although not a phoneme /k/. After the introduction of /k/ the Samoan system was as follows:

- (5) Phonemes /p t k m n g/
 | N | N
 | | | |
 Phonetic realisation [p t k m n g]

Coll. Sam. then simplified this system following the same principles underlying (3b) and (4b):

- (6) Phonemes /p k m g/
 | | | |
 | | | |
 Phonetic realisation [p k m g]

But Lit. Sam. simplified by reducing the possible allophonic realisations of the phonemes to one single allophone:

- (7) Phonemes /p t k m n g/
 | | | | | |
 | | | | | |
 Phonetic realisation [p t k m n g]

And then we are back where it all started!

6. ALTERNATIVES

The analysis given above implies that the sound changes in the different Polynesian languages are independent of each other and determined only by general phonological tendencies and the inherited phonological structure of the languages in question. Yet there are a few data that may indicate a stronger connection between the languages involved than I have till now assumed.

Tokelauan spoken just north of Samoa has retained the PP system of plosive and nasal phonemes (i.e. /p t k m n g/). But in Hale 1846 we find some curious information concerning the Tokelauan dialect of Fakafo:

The confusion in the pronunciation of k and t is not uncommon, even in those languages in which both the sounds are met with as distinct elements. In Fakafo aliti was heard for aliki and in New Zealand and Paumotu ariti. In Hawaiian, the natives make no distinction between the t and k, and the missionaries have adopted the latter, though improperly (as the element is really the Polynesian t), in the written language. (pp.233-234)

It is one peculiarity of this dialect (i.e. Fakafo - E.H.) that the k at the beginning of many words is often dropped, apparently at the mere pleasure of the speaker. (p.258)

No descriptions of Modern Tokelauan mention these phenomena and in my own fieldwork on Atafu I have not observed any cases of k-t alternation or of the

dropping of initial k's (except in recent loanwords from Samoan). But there are a few examples of g instead of n both in Tokelauan (e.g. kogā/konā *because*) and Anuta (cf. e.g. the personal name Po Tingirau in Feinberg 1982 = Samoan Tigilau, Tongan Sinilau, Rarotongan tinirau, etc.). The most likely explanation of such data and the information provided by Hale (1846) (cf. above) is to assume a certain influence or borrowings from Samoan in Anuta and especially Tokelauan.

But seen in connection with the regular development $t > k$ and $n > g$ in Luangiua³⁰ we cannot completely exclude the possibility that we once had a Samoan-Luangiua Sprachbund which also partly influenced Tokelauan and maybe Anuta and which was characterised by a phonetic instability of dental and velar plosives and nasals with a large degree of phonetic overlapping in the phonemes concerned. But only further research on the Polynesian outliers and Tokelauan can clarify this hypothesis.

NOTES

1. This work is partly based on field studies supported by The Institute for Comparative Research in Human Culture, Oslo and The Scandinavian Institute of Asian Studies, Copenhagen. I wish to thank Niko Besnier and Sheldon Harrison for comments on specific points and especially Marianne Haslev Skånland for detailed comments on an earlier draft of this study. I am also grateful to the Kon-Tiki museum, Oslo, for giving me access to its excellent Polynesian library and to the School of Oriental and African Studies for giving me access to the archives of the London Missionary Society and the Methodist Missionary Society. All quotations below from the journals of the LMS missionaries and Methodist missionaries are based on material from those archives.
2. Except in quotations, the velar nasal will in this paper be written g conforming to the Samoan orthography.
3. PP = Proto-Polynesian.
4. In all descriptions of Samoan this phoneme is called a glottal stop or break. But in Mod. Sam. it is only realised as a glottal stop in phrase-initial position. In intervocalic position it is mainly realised as creaky voice and only sporadically as a glottal stop.
5. Mod. Sam. = Modern Samoan, i.e. the t/n-style as used today in speaking and writing, cf. Coll. Sam. = Colloquial Samoan, i.e. the k/g-style and Lit. Sam. = Literary Samoan, i.e. the written language of the Samoan Bible, etc.
6. Attested as cochì in Martin 1817, II.
7. The first missionaries (through most of them having a Tahitian background where a similar development $*k > '$ had occurred) also observed the correspondence Sam. ' = k in many other Polynesian languages, cf. e.g. A. Buzacott's journal from 1836/37 where we find the following note: "The true Samoan has neither the k nor the h mostly where those letters occur in the corresponding words of the Tahitian & Rarotongan dialects, the Samoan substitutes the f for the h, & a kind of break for the k."

8. The word is attested as early as 1793 in Tongan in the form *oucaméa fer*, cf. Rossel 1808:562, cf. also *ookumméa metal* in Martin 1817, II.
9. cf. *goolo pot* in Martin 1817, II.
10. Shore 1982:268 assumes that *saka* predates European contact, but this assumption is rather questionable, cf. Love 1983:142 and note 15 below.
11. cf. Shore 1982:267-283 for a survey of the social aspects of the two sociolects, their status and range of use and the rather strong emotional attitudes manifested towards improper use of them.
12. This may be a borrowing from Tongan, cf. Tongan *puke to take hold of, seize, make a grab at*, being used in a slang-like way in Samoan, cf. above.
13. A curious exception is *tītata*, *tīkata*, *kītata tea-kettle*, borrowed from English but probably via Tahitian (cf. Tahitian *tītata*). Let me also add that once during a Sunday service I heard the form *lesu Teriso* (= *lesu Keriso*).
14. Some of the hypercorrect forms have become standard in Mod. Sam. like e.g. *toniga uniform*, cf. *tōgiga* in Pratt 1977 and Milner 1966. Even in Lit. Sam. we find a few hypercorrect forms like *fesoasoani help* (< **fesoasoagi*, cf. Milner 1966:212) and *fealofani love each other* which should have been **fealofagi*, cf. *alofagia*. But as Niko Besnier has pointed out to me, some of these variations may be due to the productivity of the *-gia* suffix in Samoan, Tuvaluan and maybe other Polynesian languages of that area.
15. It is interesting to observe that the following note is added to the interjection *puke* in Pratt 1862:170: "It and the following compound (= *puketā*) are the only instances in which the *k* is used". In later editions of the dictionary (cf. Pratt 1977:236) he has made a significant change: "It and the next word were the only instances in which the *k* was used until the recent corruption of *t* into *k*."
16. Here quoted from the 4th edition of 1893.
17. Buse ends his note by quoting a "footnote added to the second edition of Pratt's *Samoan grammar and dictionary* (1876)" which is identical (except for a few minor details) to the quotation from Whitmee quoted by Von Bülow, above.
18. None seems explicitly to regard them as separate, but many mention only the change *t > k*, while *n > g* tends to pass unnoticed. On the other hand it is worthwhile to observe that Hale 1846 did not mention the change of plosives at all, but was fully aware of a certain instability of the nasals, cf. the following note.
19. cf. also Stair 1897:15-16: "Even the Samoan language is changing. I was lately speaking to a young Samoan, a Malua student too, and could hardly understand him, as he turned all the *t*'s into *k*'s". Stair lived in Samoa from 1838 to 1845 and his remark must imply that the change had not taken place or was quite uncommon in that period. It is also interesting to note that Hale 1846 does not mention *t > k* in Samoan at all, and seems to consider *n > g* as a sporadic change mainly representing a kind of assimilation or metathesis (Hale 1846:234), cf. e.g. *manogi > magogi/magoni* and *manutagi > magutagi* as well as the example *lagona > lagoga* quoted from Pratt 1862, cf. above.

20. cf. also Von Bülow 1897b:345, Smith 1898:141, and Hocart 1916:42. Violette 1870:208 states that k "paraît naturel aux habitants de la partie sud-est d'Upolu". An interesting, although perhaps not quite trustworthy socio-linguistic observation is found in Wendt 1977:50 where he tells the story about some young men visiting Apia for the first time (1921). They found "that the Samoan spoken by the inhabitants of Apia was quaint, unusual - they used the k instead of the t. Perhaps it had something to do with the language of the papalagi, surmised Osovae". Some authors are less precise, e.g. Churchward 1926:16 who talks about these changes having taken place "in most parts of the group" without further attempts to make the areas explicit.
21. There are not many who explicitly support this hypothesis, but curiously enough none who explicitly argue against it.
22. Shore 1982:271 lists as some of the most typical contexts for the formal pronunciation (= t/n-style) schools, churches, radio broadcasts, conversation with palagis, reading and writing.
23. Quotation from Milner 1966:xiv omitted (E.H.).
24. The book is called *E TALA A, E, F*, contains 11 pages and consists of spelling lessons and 14 short chapters of reading lessons. The book was printed at the Mission Press, Huahine 1834. No author is mentioned on the front page, but there is no doubt about the author, cf. G. Pratt's journal from 6 September 1835 and Barff's account of Buzacott's labours (written 1847/48, but referring to his and Buzacott's visit to Samoa in 1834).
25. cf. the following extract from G. Pratt's journal of 14 October 1835:
- In the evening with the help of the native teachers and some boys who also know Tahitian transposed a Hymn that they might at least sing with understanding as they have hitherto learned to read and to sing in a foreign language, tho the teachers give the sense of the Hymn in the language they always sing it in the Tahitian language which only a few who can read know.
26. cf. the following note from Hardie's journal of 1838:
- Found the settlement (in Palauli) althogether in a very unpromising state: the school being badly attended and but very little attention paid to divine things. The Tahitian teacher, who has been here for some time is very inefficient, being advanced in life and having acquired but a very inadequate knowledge of the language.
27. Unfortunately we know very little of the Samoan literary language established by the Wesleyan missionaries from Tonga. None of the books they published in Samoan have survived as far as I know, but judging from the words and sentences in Samoan in the letters and journals of Peter Turner, they seem to have chosen a norm quite similar to the one the LMS chose. Having Tongan as their starting point, this may be explained in the same way as the LMS missionaries' choice of a norm for the language. But I hesitate to draw any definite conclusions here since there are hints in Turner's journals that k was not a marginal letter in his alphabet, cf. his journal of 29 January and 21 May 1835.

28. cf. the quotations from the influential Methodist missionary George Brown (Brown 1916) given above.
29. /l/ is not included in (2), being outside the labial/dental/velar system of sounds.
30. Another problem is that we do not know the exact chronology of these sound changes in Luangiua. Brown 1910:414 gives the name of the island as Lua Niua which seems to indicate that the change $n > g$ is rather recent there. But until further material is available we cannot exclude the possibility that Brown was influenced by his Samoan background and invented a hyper-correct Luangiua form.

REFERENCES

BEAGLEHOLE, J.C.

- 1963 *The Endeavour journal of Joseph Banks 1768-1771*, I-II. Sydney: Angus and Robertson.
- 1967 *The journals of Captain James Cook on his voyages of discovery: the voyages of the Resolution and Discovery 1776-1780*, I-II. Published for the Hakluyt Society. Cambridge: At the University Press.

BIGGS, Bruce

- 1978 The history of Polynesian phonology. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian linguistics: proceedings: Fascicle 2:691-716*. PL, C-61.

BROWN, George

- 1910 *Melanesians and Polynesians: their lifehistories described and compared*. London: Macmillan.
- 1916 Remarks on Samoan sound changes. *Man* 16:181-183.

BROWNE, Arthur H.

- 1897 An account of some early ancestors of Rarotonga. *Journal of the Polynesian Society* 6:1-10.

BUCK, P.H. (Te Rangi Hiroa)

- 1930 *Samoan material culture*. Honolulu: B.P. Bishop Museum Bulletin 75.

BÜLOW, W. von

- 1897a Samoanische Schöpfungssage und Urgeschichte. *Globus* 71:375-379.
- 1897b Die samoanische Sprache von Heute. *Zeitschrift für afrikanische und oceanische Sprachen. Mit besonderer Berücksichtigung der Deutschen Kolonien* 3:344-352.
- 1900 Die Namen der Samoa-Inseln. *Globus* 78:31-33.

BUSE, J.E.

- 1961 Two Samoan ceremonial speeches. *Bulletin of the School of Oriental and African Studies* 29:104-115.

CHURCHWARD, Spencer

1926 *A new Samoan grammar*. Melbourne: Spectator Publishing Co.

ELBERT, Samuel H.

1982 Lexical diffusion in Polynesia and the Marquesan-Hawaiian relationship. *Journal of the Polynesian Society* 91:499-517.

ELLIS, W.

1831 *A vindication of the South Sea Missions from the misrepresentations of Otto von Kotzebue*. London: Friedrich Westley and A.H. Davis.

FEINBERG, Richard

1982 Structural dimensions of sociopolitical change on Anuta. *Pacific Studies* 5/2:1-11.

FUNK, B.

1893 *Kurze Anleitung zum Verständnis der samoanischen Sprache: Grammatik und Vokabularium*. Berlin.

GILL, William Wyatt

1876 *Myths and songs from the South Pacific*. London: Henry S. King.

HALE, Horatio

1846 *Ethnography and philology. United States exploring expedition during the years 1838, 1839, 1840, 1841, 1842 under the command of Charles Wilkes*. Philadelphia: Lea and Blanchard.

HOCART, A.M.

1916 A Samoan sound change. *Man* 16:42-43.

KOTZEBUE, Otto von

1830 *A new voyage round the world, I-II*. London: Henry Colburn and Richard Bentley.

KRÄMER, Augustin

1897 Bericht über neue samoanische Überlieferungen. *Globus* 712:76-78.

1903 *Die Samoa-Inseln, II*. Stuttgart: E-Schweizerbartsche Verlagsbuchhandlung.

LOVE, Jacob Wainright

1983 Review of Shore 1982. *Pacific Studies* 7:122-145.

LUNDIE, G.A.

1846 *Missionary life in Samoa, as exhibited in the journals of the late G.A. Lundie, during the revival in Tutuila in 1840-41*. Edinburgh: Oliphant.

MARTIN, John

1817 *An account of the natives of the Tonga Islands, I-II*. London: John Murray.

MILET-MUREAU, L.A., ed.

1797 *Voyage de La Pérouse, I-IV*. Paris: L'Imprimerie de la République.

MILNER, George B.

1966 *Samoan dictionary*. London: Oxford University Press.

MOYLE, Richard

1981 *Fāgogo: fables from Samoa in Samoan and English*. Auckland: University of Auckland Press.

NEFFGEN, H.

1904 *Deutsch-samoanisches Konversationsbuch. O le Tusi-Fetalia'i fa'asiamani ma fa'asamoa*. Leipzig: Otto Fischer Verlag.

PRATT, George

1862 *A Samoan dictionary: English and Samoan and Samoan and English with a short grammar of the Samoan language*. Samoa: London Missionary Society Press.1977 *Pratt's grammar & dictionary of the Samoan language*. 4th edn, 1911, reprinted. Apia: Malua Printing Press.

ROSSEL, E.P.E. de, ed.

1908 *Voyage de D'Entrecasteaux, envoyé à la recherche de La Pérouse ...*, I. Paris: L'imprimerie impériale.

SHORE, Bradd

1982 *Sala'ilua: a Samoan mystery*. New York: Columbia University Press.

SIERICH, O.

1900-04 Samoanische Märchen. *Internationales Archiv für Ethnographie* 13:223-237, 14:15-23, 15:167-200, 16:88-110, 17:182-188.

SMITH, S. Percy

1898 *Hawaiki: the whence of the maori: being an introduction to Rarotonga history*. *Journal of the Polynesian Society* 7:136-177.

STAIR, John B.

1897 *Old Samoa, or flotsam and jetsam from the Pacific Ocean*. London: Religious Tract Society.

STUEBEL, O.

1895 *Samoanische Texte*. Herausgegeben von F.W.K. Muller. Königliche Museen zu Berlin. Veröffentlichungen aus dem königlichen Museum für Volkerkunde IV. Baand - 2.-4. Heft. Geographische Verlagsbuchhandlung Dietrich Reimer. Berlin.

THOMSON, Basil, ed.

1915 *Voyage of H.M.S. Pandora*. London: Francis Edwards.

VIOLETTE, Le R.P.L.

1870 *Notes d'un missionnaire de Samoa (Océanie centrale). Les Missions catholiques. Bulletin hebdomadaire de l'oeuvre de la propagation de la foi*, III. Lyon.

WALPOLE, Fred

1849 *Four years in the Pacific in Her Majesty's Ship "Collingwood" from 1844 to 1848*, II. London: Richard Bentley.

WENDT, Albert

1977 *Pouliuli*. Honolulu: The University Press of Hawaii.

WILLIAMS, John

1834 *E TALA A,E,F*. Huahine: Mission Press.

1837 *A narrative of missionary enterprises in the South Sea Islands*.
London: J. Snow.

LINGUISTIC CONVERGENCE IN CENTRAL VANUATU

Ross Clark

Two Polynesian Outlier languages, Emae and Mele-Fila, are spoken within about 80 km of each other in central Vanuatu (see Map). Whereas the majority of Outliers in the north are spoken on truly isolated islands, the Emae and Mele-Fila speakers are nowhere more than about an hour's travel (on foot or by canoe) from speakers of non-Polynesian languages. The close cultural ties among the peoples of this region have been well documented. Allen (1981:5) describes Efate and the small islands to the north as "a single cultural area characterized by the presence of dispersed matrilineal clans and an hereditary titular system". Guiart (1973) provides abundant evidence of the similarities in social organisation and the complex network of political and mythological connections. Other writers (Nevermann 1953, Simmons et al 1954) have observed the lack of any clear physical difference between the Polynesian speakers and their neighbours. All this indicates a long period of physical and cultural assimilation. The present paper deals with the results of the corresponding process of convergence in language.

The languages in question are:

EMAE (Polynesian), 200 speakers,¹ in Tongamea and Makatea villages at the eastern end of Emwae Island;

MELE-FILA (Polynesian), 1800 speakers, in Mele (Imere) village, north-west of Vila, and on Fila (Ifira) Island, at the entrance to Vila Harbour. Mele-Fila and Emae both belong to the Samoic-Outlier subgroup of Polynesian (Pawley 1967, Clark 1978), but do not appear to be closely related beyond that. They share between 40 and 50 percent cognates on the Swadesh 200-word list.²

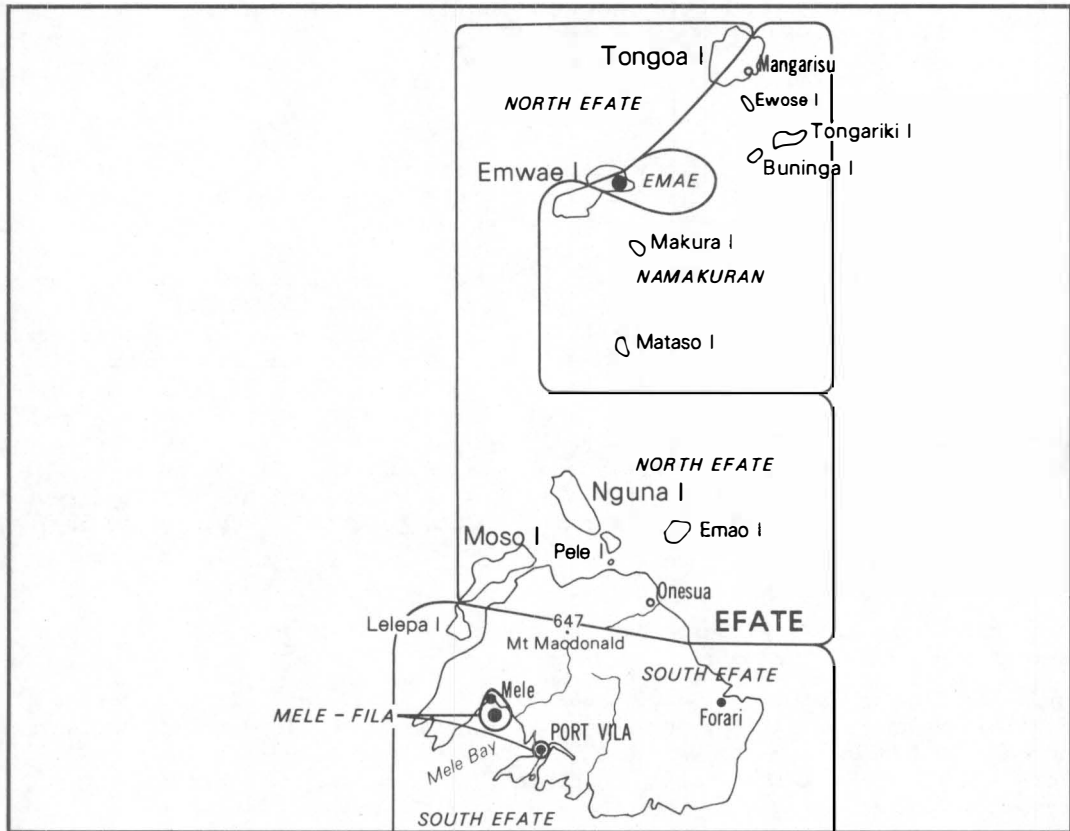
NAMAKURA (Melanesian),³ 2000 speakers, on Makura, Mataso, Buninga, Tongariki and part of Tongoa, as well as at Sangava and nearby settlements on the south coast of Emwae;

EFATE (Melanesian), 4700 speakers, on Efate and its offshore islands, part of Tongoa, and in Sesake and Marae villages on the north side of Emwae. Tryon (1976, 1981) treats Efate as two languages, but I prefer to consider it as a single dialect chain. Namakura and Efate are linguistically next-of-kin, sharing 50-60 percent cognates in basic vocabulary. They make up a distinct subgroup within the North and Central Vanuatu subgroup of Oceanic (Tryon 1976, Clark 1985).

I will first consider the influence of the Melanesian languages on the Polynesian, which has been extensive in all areas. I then turn to the much more restricted Polynesian influence on Melanesian, and finally consider explanations for the observed convergence effects.

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 333-342. *Pacific Linguistics*, C-94, 1986.

© Ross Clark



Map 1: Central Vanuatu

1. MELANESIAN INFLUENCE ON POLYNESIAN LANGUAGES

1.1 Lexical

A sample of general vocabulary in Mele-Fila and Emae shows the following origins:

	MELE-FILA	EMAE
Polynesian	36 percent	52
Melanesian	20	12
Equivocal	12	10
Other	6	8
Unknown	26	18

The 'Polynesian' and 'Melanesian' categories include only words for which clear etymologies are known. 'Equivocal' items are those for which PN and MN forms converge (e.g. MF *Em manu bird*, cf. PPN **manu*, Ef *manu*), or which contain two morphemes of different origins (e.g. MF *vaisara stream*, from PPN **wai water*, Ef *sara flow*). 'Other' consists of modern borrowings from English, French and Samoan. The 'Unknown' component can be expected, with further research, to be reassigned mainly to 'Melanesian' rather than 'Polynesian'.

Mele-Fila obviously has a stronger lexical component of Melanesian origin than does Emae; the latter at once gives one the impression of being 'more Polynesian'. The picture does not change markedly if we restrict our attention to core vocabulary. The following figures are based on the 292-item list used in Tryon 1976:

	MELE-FILA	EMAE
Polynesian	56	70
Melanesian	24	7
Equivocal	10	10
Other	- ⁴	0
Unknown	10	13

Contrary to normal expectations that core vocabulary should be relatively imper-vious to borrowing, the Melanesian percentage in Mele-Fila actually appears to increase. This is probably an accidental effect resulting from the smaller number of 'Unknown' items. If we compare just the unequivocal Polynesian and Melanesian items, the expected decrease appears. Nevertheless, in both languages the Melanesian component is of the same order of magnitude in core as in general vocabulary.

Lexical influence without borrowing of actual forms, i.e. calquing or re-modelling of semantic fields, is less easy to detect, but seems to be fairly common. For example, MF *nifo tooth, seed* is formally derived from PPN **nifo tooth*, but has apparently acquired its additional sense from Ef *na-pati tooth, seed*. Emae *matua old man, husband* combines the form of PPN **matu'a elder, parent* with the meaning of Ef *maariki, Nmk ma'arik*. For the same semantic field, Mele-Fila has *nuaane* - apparently an analogically created counterpart to *nuufine old woman, wife*, from PPN **nu(a)fine old woman*. The PPN form **afaa hurricane* is replaced by MF *matagi taa*, evidently modelled on Ef *na-lagi atu*, both analysable as *wind + hit*(intr.). Adjustments of this sort tend towards a state in which the different languages are simply, as Grace (1981) puts it, different "lexifications" of the same set of fields.

1.2 Phonological

There is little scope for convergence in the vowel system, since all four languages have the same set of five vowel qualities /i e a o u/. Length appears to be contrastive in all four as well, though with a lower functional load in Melanesian than in Polynesian.

A much clearer contrast appears in the consonants. Table 1 compares the Polynesian consonant inventory (common to Emae and Mele-Fila) with those of Efate and Namakura. It will be seen that the Melanesian systems have additional contrasts at almost every point - the one exception being the contrast of voice in the labial fricatives, which is found only in Polynesian.

The result of contact has been an expansion of the consonant inventory of both Polynesian languages:

- (1) Both Mele-Fila and Emae have added /l/ and /w/.
- (2) Mele (but not Fila) has added the labiovelars /β/ and /ɱ/.

(3) Emae has added the prenasalised stops /b/, /d/ and /q/. The fact that Mele-Fila, in general much more Melanesian-influenced than Emae, has not acquired these consonants can be explained by the fact that the southern Efate dialects,

with which Mele is likely to have been in closest contact, do not have them. Mele does, however, have a small number of words with /n/ for Efate /d/ (MF *nina*, Ef *liida* *wasp*, with assimilation; MF *peana*, Ef *na-peada* *arrowroot starch*; MF *tāmaaniku*, Ef *tāmaduku* *chiton*). These certainly suggest that the immediate source had a prenasalised /d/, which was assimilated to the nearest Polynesian consonant rather than incorporated as such.

Table 1: Consonant inventories compared

POLYNESIAN	EFATE	NAMAKURA
p	p	p
	ḽ	ḽ
t	t	t
k	k	k
	b ¹	b
	ḽ ¹	ḽ
	d ¹	d
	q ¹	q
f	f ²	
v	v ²	v
s	s	s
		h
m	m	m
	ḿ	ḿ
n	n	n
g	g	g
	l	l
r	r	r
	nr ³	
	w	w

Notes:

1. Prenasalised stops occur only in the northern dialects of Efate.
2. Northern Efate dialects have v, southern dialects have f, but no dialect contrasts the two.
3. Southern dialects only.

In the orthography used in this paper, b, d, and q are prenasalised stops; nr is a prenasalised trill; g is a velar nasal; ḽ, ḽ, and ḿ are labio-velar consonants. Efate forms cited are in the Nguna (North Efate) dialect, unless otherwise indicated.

For the most part, the new consonants just mentioned are found only in borrowed vocabulary (including more recent loanwords from English, French and Samoan), and all borrowed items retain these consonants as in the source language. A few exceptions to both parts of this statement, however, are worth noting. A small number of words in Mele-Fila show MF /r/ for Efate /l/ and plain labials for labio-velars: the names of the two islands, MF *l-mere*, *l-fira*, Ef *E-ḿele*, *E-fila*, and the homophonous word for the cycas palm, MF *mere*, Ef *na-ḿele*; MF *panu*, Ef *na-ḽanu* *mat*; MF *pokasi*, Ef *na-ḽokasi* *pig*.⁵ It seems most likely that these were among the earliest borrowings, at a period when the new consonants

had not been established in MF, and that they were too common and well-established to be re-shaped or re-borrowed at a later date. Their semantic distribution, too - place names, objects of trade and ritual - suggests an early stage when contact was not as intimate as it was to become.

An equally small group of Emae words, which must be of Polynesian origin, shows consonants which normally occur only in the non-Polynesian vocabulary (Em lasi *big*, from PPN *lasi, where PPN *l normally becomes Em r; Em bakakau *wing*, from PPN *pakakau, where PPN *p normally remains p in Emae). This phenomenon of phonological spill-over as a result of borrowing does not seem to be widely recognised or discussed, though Anttila (1972:168) gives a few examples, and suggests hypercorrection and emotional associations as possible explanatory factors.

A phonological innovation which clearly has spread beyond borrowed words is Mele-Fila stress, which is now on the antepenultimate vowel (as in Efate), rather than the penultimate (as in most Polynesian languages, including Emae).

1.3 Grammatical

Both Emae and Mele-Fila have been extensively influenced grammatically by the Melanesian languages. The following is a summary of some of the major changes.

(1) In major constituent order, both Emae and Mele-Fila are strictly SVO, as are Efate and Namakura. In other Polynesian languages, verb-initial order is either obligatory or at least a normal alternative.

(2) Most Polynesian languages use a variety of verbless predicate structures to assert identity, location and possession (Clark 1969). The corresponding predicates in Emae and Mele-Fila, as in Efate, use lexical verbs which can be glossed *be* or *have*. The forms of these verbs and their origins are shown in Table 2.

Table 2: Verbs for 'have' and 'be'

	<i>be (something)</i>	<i>be (in a place)</i>	<i>have</i>
Efate	vei	toko	veani
Mele-Fila	fei (< Ef)	tuu (< PPN *tu?u <i>stand</i>) lawo (? Ef lawo <i>stand</i>) tokoto (< PPN *takoto <i>lie</i>) etc.	lekina (< Ef vei talakea-na <i>be its owner</i>)
Emae	tupu (< PPN *tupu <i>grow</i>)	tuu nofo (< PPN *nofo <i>sit</i>) etc.?	tuni (< ?)

Note: The different verbs of location in Mele-Fila and Emae are selected by the nature of the subject: *tuu* for persons, *lawo* for large objects, *tokoto* for small objects, etc.

(3) A number of connectives and subordinate clause markers in Mele-Fila and Emae are lexically derived in a way that mirrors their Efate counterparts:

	EFATE	MELE-FILA	EMAE
<i>when</i> (adverbial) = (that) time	(ragi) waina	napoo	turaga
cf. <i>now</i> = <i>this time</i>	ragi waia	napoo naa	turaga ni
<i>because</i> = <i>its base</i> (of it)	na-lake-na	ton-lake (gani)	na tafito (ai)
<i>until</i> = <i>go go</i>	paa-paa	fan-fan	ano-ano

(4) A conspicuous feature of all of these languages is the use of compound verbs. These number in the hundreds in my data, and the process of compounding seems to be productive to some extent. Compound verbs are typically transitive, and the second verb bears the transitive suffix. In Emae the first verb may also be suffixed, but in Efate and Mele-Fila it may not. In Mele-Fila the first verb may have a morphophonemically reduced form, as in *man-saraavia forget*, from *mantua* (PPN **manatu*) *think, remember* + *saraavia miss* (PPN **sala error*). As this example shows, compounding in the Polynesian languages is not restricted to borrowed verbs. A further example will illustrate this for Emae, as well as further exemplifying the greater infiltration of actual borrowed lexical items in Mele-Fila. The original model is Efate *vasa-ṗotae explain, judge* (= *speak* + *divide*). Mele-Fila *pasa-wotaaea* is constructed with the same actual forms, whereas Emae *muna-vaea* uses verbs of Polynesian origin with the same meanings (PNP **muna speak* (*confidentially*), PPN **wahe divide*).

(5) The Polynesian possessive system historically involved two possessive categories marked by *o* and *a*. In addition to what is commonly considered 'inalienable' possession (part-whole and kin relations), the PPN **o* category included certain important material possessions. (See Wilson 1982, Chapter 2, for a full discussion.) In Mele-Fila, the possessive category marked by what are historically *o*-forms is now a more restricted 'inalienable' category agreeing with the Efate category marked by suffixed possessives. In addition, the formal symmetry between *a* and *o* forms which is normal in Polynesian languages has been completely lost, and the historical *a*-forms are now postposed to the noun, like their Efate counterparts:

	<i>my leg</i>	<i>my canoe</i>	<i>my fowl</i>
PPN	*tooku wa'e	*tooku waka	*ta'aku moa
Mele-Fila	tuku-vae	te-paki neaku	te-moa neaku
Efate	na-tua-gu	rarua aginau	tooa aginau

(6) Mele-Fila has lost the productive use of the Polynesian locative preposition *i*, which now occurs only as a prefix on a closed class of locationals (MF *i-ruga above*, *l-mere Mele*, cf. Efate *e-lagi*, *E-mele*). Other locative phrases are unmarked, as in Efate and also in Emae:

Mele-Fila	rat e-nnofo te-panu
Efate	eu taasake na-ṗanu
Emae	tere nofo re bele

They sat on the mat

(7) Two other PPN case markers, *ki (goal) and *ma (comitative) are replaced in Mele-Fila by gaia and soina respectively.⁶ These appear to be historically transitive verb forms (PPN *haŋa-ia *face towards, turn towards* and *soa-ina *be companion/partner of*), and still have some verbal properties: soina can be used as an independent verb (ki to soina maateu *you will be with us*), and gaia can be followed by certain post-verbal particles. Their Efate counterparts are likewise verbal: paki *go to* marks goal phrases, and sikoti *accompany* marks comitatives.

(8) Mele-Fila, Emae and Efate all have a distinction of two 'future' categories, formally expressed in parallel ways. One preverbal particle by itself expresses immediate intention, or an imperative/hortative sense. When followed by a second particle, the meaning is a more general or remoter future.

	<i>let's go</i>	<i>we will go</i>
Mele-Fila	tu tee-roro	tu too-roro
Emae	tu ka ano	tu ka po ano
Efate	tu ga vano	tu ga wo pano

MF tee and Em ka are formally of Polynesian origin (PPN *te 'non-past', *ka 'anticipatory'), but MF too and Em po are not.⁷ If we assume that too < *te + wo, it seems at least possible that in both cases the second particle is borrowed from Efate, though the Emae consonant is unexpected.

(9) Another apparent example of borrowing of a verbal particle is the Emae conditional marker pe: Em kere pe ano *if you go*, cf. Ef ku pe vano. Although this may be connected with PPN *pe *or, whether*, its exact syntactic and semantic properties agree better with Efate than with any other Polynesian language.

2. POLYNESIAN INFLUENCE ON MELANESIAN LANGUAGES

The idea of 'Polynesian influence' in Melanesia has a long and not entirely happy history. The whole question is still haunted by 19th century racial assumptions, under which any trace of lighter skin, hereditary chieftainship or straightforward sound correspondences in Melanesia was interpreted as indicative of 'Polynesian admixture' into a basically quite distinct Melanesian race, culture or language. Perhaps the development of a clearer idea of Melanesian-Polynesian relations, and the accumulation of better descriptive material, in recent decades, may now make possible a fresh assessment of the situation.

In the present case, all I am able to show here is a list of words from Nguna, a northern Efate dialect (Table 3) which I believe have a good chance of being Polynesian borrowings. The criteria for this judgment are various, and are not set out in detail here, but they are basically of two sorts: distributional and phonological. A word is considered a likely PN loan if it is widely distributed in PN languages outside Vanuatu, but not in NCV languages; or, if it shows sound changes which would not be expected in Nguna but could be accounted for by inheritance via a PN language. An instance of the latter is Nguna voonu *turtle*, from PO *poñu. As Nguna regularly shows w for PO *p before o, this form cannot be directly inherited. However, a reflex of PPN *fonu would be borrowed by Nguna in just this form. Here the conclusion of borrowing is supported by distributional evidence, since Nguna is the only NCV language which has a reflex of the Proto-Oceanic form.

Table 3: Probable Polynesian loanwords in Efate (Nguna)

peao	<i>wave on the open sea</i> (PPN *peau, Em piau, MF peau)
na-raki	<i>cold wind</i> (PPN *laki <i>west wind</i>)
rarua	<i>canoe</i> (PPN *laa rua <i>two sails</i>) cf. Futuna-Aniwa rarua <i>two-masted canoe</i>
ta(v)ura	<i>whale</i> (PPN *taf(o,u)raʔa, Em MF tafuraa)
na-tira	<i>mast</i> (PPN *tila, Em tira, MF jira)
vonu	<i>turtle</i> (PPN *fonu, Em MF fonu)
vorau	<i>to sail</i> (PPN *folau, Em MF forau <i>travel by canoe</i>) cf. directly inherited Ef wowolau <i>steer a canoe</i>
kaaka	<i>fibre at base of coconut frond</i> (PPN *kaka, Em kaka, MF mukaka)
kavekave	<i>kind of basket</i> (MF kavekave, cf. PPN *kawe <i>carry</i>)
kokovu	<i>cook food wrapped in leaves</i> (PPN *kofu, MF kofu-a)
koovu	<i>laplap with fish inside</i>
na-moega	<i>kind of mat</i> (PPN *mohega <i>sleeping-mat</i> , MF moega)
na-pora	<i>woven coconut-leaf wall panel</i> (PPN *pola, Em MF pora) cf. directly inherited Ef polo <i>coconut leaf basket</i>
rekei	<i>to decorate</i> (PNP *laakei, MF fakarakeia)
na-roro	<i>coconut oil</i>
roroi	<i>to squeeze coconut cream onto</i> (PPN *lolo-ʔi, Em roro-i, MF rroi)
na-tara	<i>bottom end of thatched roof</i> (PPN *tala <i>end of house</i> , MF tara <i>eaves</i>)
mori	<i>to accompany, escort person to a place; to pay</i> (PPN *mori, Em MF mooria <i>escort</i> , MF mmori <i>wages, pay</i>)
sosori, soria	(Em MF soria)
vakotovi	<i>buy, pay</i> (PSO *fakatau-ia, Em fakatauvia, MF fakatawia)
vono	<i>talk, discuss</i> (PPN *fono)
voroa-ki	<i>assign inheritance or succession</i> (PPN *poro(aki) <i>give (parting) instructions</i> , Em poroaki-na <i>say goodbye</i>)
na-tipua	<i>kind of spirit, dwarf</i> (PPN *tupuʔa, MF tupua <i>supernatural being</i>)

What is striking about the list in Table 3 is first its small size - only a couple of dozen borrowings out of 1500-2000 lexical items in my file⁸ - and second its semantic concentration in certain areas: the sea and navigation, material culture, politics and trade. This is almost archetypal 'cultural borrowing', and even fits the traditional idea of Polynesians as more expert seafarers than the Melanesians.

It would be premature to say that there has been no phonological, grammatical or semantic influence of Polynesian on Melanesian languages in Central Vanuatu. Our reconstructed baseline for the Melanesian languages is not nearly so clearly drawn as PPN. It may be that influences in these areas will yet be detected. But they will certainly be of a relatively minor and subtle sort, compared with the rather drastic effects of Melanesian on Polynesian.

3. DISCUSSION

How are we to account for the asymmetry in the convergence effects observed in the two directions we have been considering? To do so we must make a clear distinction between 'cultural' and 'intimate' borrowing.⁹ Cultural borrowing requires only some type of contact between the two speech communities. It is basically lexical, and is concentrated in areas where the two cultures differ (including flora and fauna in situations where one group is immigrant). Intimate

borrowing requires that second-language speakers (trade partners, in-marrying spouses, etc.) play a major part in the life of the borrowing community. It affects all areas of the language at once.

We have seen (Table 3) clear evidence of cultural borrowing from Polynesian into Melanesian. Almost certainly cultural borrowing also took place in the opposite direction, but its effects have been all but swamped by the massive intimate borrowing from Melanesian into Polynesian. However, as noted above, a small group of words suggestive of the earliest period of contact (on the basis of deviant sound correspondences) are in fact semantically concentrated in typical cultural-borrowing areas.

The reason why Melanesian shows no signs of intimate borrowing from Polynesian in this case is probably to be explained simply on numerical grounds. We can safely assume that the first Polynesian-speaking immigrants were few in number and found an established Melanesian population. Even today, there are more than three times as many Melanesian as Polynesian speakers in the Efate region. If, as this suggests, Polynesians have always been a minority, they would, in establishing trade contacts or seeking spouses outside the village, have had to deal with Melanesian speakers more often than not, whereas Melanesian speakers, on average, would have had only a minority of Polynesian contacts. Melanesian wives, in particular, marrying into Polynesian villages, bearing and rearing children, speaking a Melanesian-influenced second-language variety, would have accomplished both the physical and the linguistic assimilation of the immigrants.

NOTES

1. The numbers of speakers are as given by Tryon 1981.
2. This smooth phrase actually covers two rather discrepant computations. 38 percent by Peter Ranby (personal communication) and 51 percent by myself. I have not looked into the reasons for this difference.
3. With apologies, I use the word 'Melanesian' to mean 'non-Polynesian Oceanic'.
4. The word *soolo* was given by my Mele informants for *salt* on the Tryon list.
5. Na-*ṽ*okasi actually appears to mean *meat* in all Efate dialects today, being replaced in the sense *pig* by Proto-Efate *waaqo. Cognates in other closely related languages, however, all mean *pig*; and the name of the mountain peak in south-west Efate - *ṽ*Pau-na-*ṽ*okasi *pig's head* - makes no sense unless the word had this sense in Efate in the not too distant past.
6. *ki survives only as an older-generation variant form with a few locational bases, e.g. *kiaro* = *gaia iaro downwards*.
7. Niuean has a future particle *to*, but to connect this with the MF form would seem to raise more problems than it solves.
8. I am indebted to A.J. Schütz and Ellen Facey for use of their unpublished lexical files.
9. I take these terms from Bloomfield (1933: chapters 25 and 26), though my definition of the two processes would not agree with his in every respect.

REFERENCES

ALLEN, Michael, ed.

- 1981 *Vanuatu: politics, economics and ritual in island Melanesia*.
Sydney: Academic Press.

ANTTILA, Raimo

- 1972 *An introduction to comparative and historical linguistics*. New
York: Macmillan.

BLOOMFIELD, Leonard

- 1933 *Language*. New York: Holt, Rinehart.

CLARK, Ross

- 1969 Some verbless sentences in Samoan. *Oceanic Linguistics* 8:105-119.
1978 The New Hebridean Outliers. In S.A. Wurm and Lois Carrington, eds
*Second International Conference on Austronesian Linguistics: pro-
ceedings*, 911-928. PL, C-61.
1985 Languages of North and Central Vanuatu: groups, chains, clusters
and waves. In Andrew Pawley and Lois Carrington, eds *Austronesian
linguistics at the 15th Pacific Science Congress*, 199-236. PL, C-88.

GRACE, George W.

- 1981 *An essay on language*. Greensboro, S.C.: Hornbeam Press.

GUIART, Jean

- 1973 Le dossier rassemblé. In J.-J. Espirat et al *Système des titres,
électifs ou héréditaires, dans les Nouvelles-Hébrides*. Paris:
Institut d'Ethnologie.

NEVERMANN, Hans

- 1953 Polynesier auf den Neuen Hebriden. *Zeitschrift für Ethnologie* 78:
196-209.

PAWLEY, Andrew

- 1967 The relationship of Polynesian Outlier languages. *Journal of the
Polynesian Society* 76:259-296.

SIMMONS, Ray T. et al

- 1955 A blood group genetical survey in New Hebrideans at Efate. *Medical
Journal of Australia* 1:547-551.

TRYON, D.T.

- 1976 *New Hebrides languages: an internal classification*. PL, C-50.
1981 Vanuatu. In S.A. Wurm and S. Hattori, eds *Language atlas of the
Pacific area*. PL, C-66.

WILSON, William H.

- 1982 *Proto-Polynesian possessive marking*. PL, B-85.

COUNTING AND CULTURE CONTACT IN NORTH-EAST NEW GUINEA

Geoff P. Smith

INTRODUCTION

This paper starts with an investigation of the numerals of the languages of the Morobe Province, Papua New Guinea. The relationship between the distribution of the different language families and types of counting system occurring in them is then examined. There is found to be a certain amount of consistency in the type of counting system found within families, but there are also unexpected variations. The limited extent of the numeral series in the Morobe Austronesian languages is particularly conspicuous, and the possible role of culture contact in this apparent regression is examined.

LANGUAGES OF MOROBE

There are just over 100 languages spoken in the Morobe Province; 40 of these are Austronesian, while 60-odd are non-Austronesian (McElhanon 1984). As with other parts of New Guinea, many of the Austronesian languages are in coastal or offshore locations, but unlike other parts of the island, there is also a considerable penetration inland. The Adzera and Buang families occupy a large inland area around Mumeng and in the Markham Valley a considerable distance from the coast at some points. In the central ranges are the non-Austronesian languages of the Angan, Binanderean and Kunimaipan families, while in the Huon Peninsula are the Eastern and Western Huon, Erap, Wantoat, Yupna and a number of minor families. It is assumed that speakers of non-Austronesian languages have been in the area considerably longer than speakers of Austronesian languages, although establishing accurate dates of settlement is difficult. The degree of penetration of the Austronesian languages into Morobe also suggests that a long interaction has taken place.

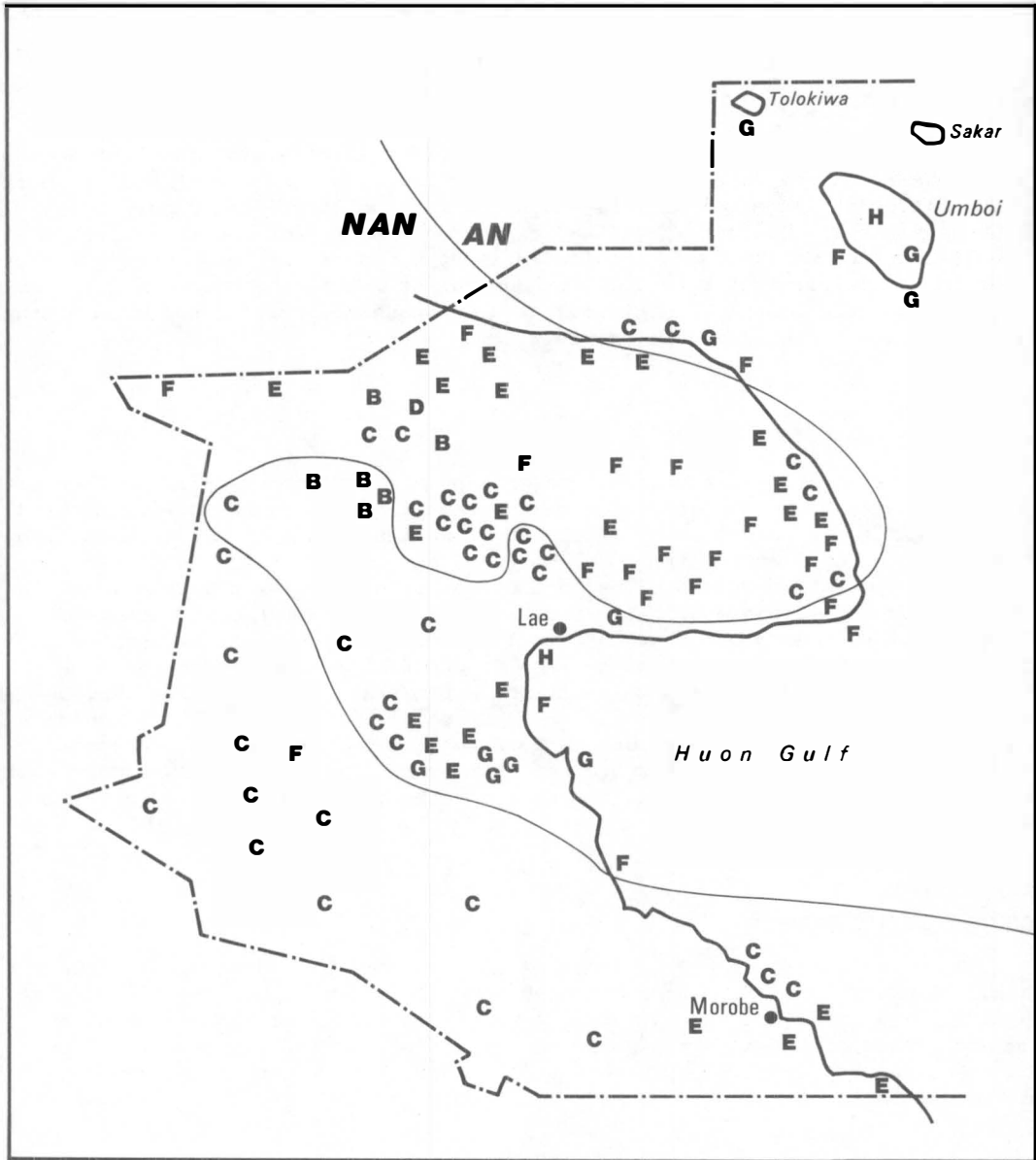
MOROBE COUNTING METHODS

An investigation of the numerals of all the Morobe languages has shown that the great majority of languages have few numerals and rely on some form of tallying for indicating higher numbers. A number of different types emerge, and the occurrence of these in the Morobe languages is indicated in Map 1. Type A, representing body counting systems found further west, is not found anywhere in the province. Types B and C have only two numeral terms, type C commonly

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 343-349. *Pacific Linguistics*, C-94, 1986.

© Geoff P. Smith

343



Map 1: Distribution of counting types in Morobe languages

using tallying on hands and feet, while type B does not. Type D has three numerals and does not refer to hands and feet, while type E has three numerals followed by tallying on the hands and feet. Type F has words for the numerals *one* to *four* followed by tallying on hands and feet to a total of 20. Type G differs in having a distinct numeral for *ten*. There may also be a distinct numeral for *five*, or the word for *hand* may be used. Type H has words for both *ten* and *20*.

There is a certain amount of consistency in the various language families: typical of the non-Austronesian families is a repertoire of two, three or four numerals followed by tallying on the hands and feet (types B to F). Austronesian languages of the Adzera and Buang families also tend to follow this pattern. Those languages with more extensive numeral series (types G and H) tend to belong to the coastal and island Austronesian languages of the Siasi family. The only exception is the non-Austronesian Kovai language of Umboi, which is surrounded by Siasi family Austronesian languages. There are, however, some Siasi family languages with very limited numeral series, for example Roinji and Nengaya, with only two numerals.

ANCESTRAL COUNTING SYSTEMS

Although the coastal and offshore Austronesian languages do tend to have more numerals than their non-Austronesian counterparts, they are deficient when compared with Proto-Oceanic and Proto-Austronesian, the inland Austronesian families remarkably so. For each of the proto-languages a complete decimal series of numerals has been reconstructed (e.g. Dahl 1976, Wurm and Wilson 1975). It seems then that in the process of culture contact, there has been a regression of the numeral series of the descendant Austronesian languages of north-east New Guinea rather than an expansion of those of the non-Austronesian languages in the area.

THE CULTURAL CONTEXT OF COUNTING

Clues to a possible explanation of this state of affairs can be gleaned from a consideration of how counting takes place in its cultural context in PNG. How do societies relying on simple tallying devices or few numerals cope with the social and economic demands of everyday life?

Recently a number of ethnographic studies have looked at this problem. In many traditional PNG societies, ceremonial exchange played an extremely important part in cultural and economic life. During many such exchanges, enumeration of totals was a necessary part of the process. However, an impressive visual display invariably accompanied presentations, and it was this which was more significant than the total itself. Many exchanges were competitive, and implied the need to give back as much as or more than what had been previously received. Thus the concept of relative quantity was usually more important than absolute measures. In these types of situations, tallying methods and poorly developed numeral series were adequate for handling all the necessary exchange procedures.

Common to many accounts of such exchanges is the importance of the concept of pattern. Large formal presentations may involve lines or heaps of wealth objects, for example laying out valued shells, or tethering pigs to stakes. Objects to be distributed may be tied to frameworks or specially prepared to maximise the impact of public display. Bunches of bananas in Adzera, for example,

were not counted prior to distribution, but mounted on a structure reaching the top of a coconut tree. The significance of this kind of structured visual display to counting behaviour was effectively summed up by Biersack (1982:30) when she stated: "relationship and hierarchy, not unit and iteration, are the sources of multiplicity".

Numbers may, in fact, sometimes appear to be redundant. Thune, working among the Loboda of Milne Bay, noted that while there exists in this society a counting system capable of reaching into the hundreds, counting and the enumeration of objects are unimportant in the culture and occur only rarely. Instead, he describes a non-numerical way of looking at the world emphasising relative rather than absolute measures. In such a situation the regression of numeral series could easily be imagined.

TRADE

If exchanges of this type were the only economic transactions existing in traditional society, the enquirer might be satisfied with this explanation. However, the Morobe Province is well known for intricate trading networks involving widely separate cultures. The trading interests of the Siassi islanders, for example, radiated like the spokes of a wheel to the far-flung corners of their traditional universe: east to New Britain, north to Umboi and Tolokiwa, west to the Madang coast and south to the Huon Gulf (Harding 1967). Each of these was the centre of trade with hinterland or other local groups. For people such as the Siassi whose life depended on trade, it would seem reasonable to assume that the possession of a complete decimal series would facilitate the trading process when, for example, calculating relative values and exchange ratios. And yet it is in just this type of cultural milieu that the regression of decimal numerals appears to have taken place. A detailed look at what took place during trading interactions may give some idea of the mathematical requirements of participating effectively.

Firstly, are we justified in using the word 'trade', with its connotations of profit, supply and demand, for the kind of ventures undertaken by the Siassi? As mentioned previously, ceremonial exchange is an important feature of Melanesian society, especially as a stage for politically ambitious 'big men', and ever since Malinowski's pioneering work on the *kula* expeditions of Milne Bay it has been recognised that trade in Melanesia also involves some of these ceremonial features. A great deal of the energy expended in *kula* shell exchanges, for example, appears to outside observers to have little justification in terms of economic benefit. Thus it might be argued that such overseas expeditions should be regarded not so much as trading ventures as complex social rituals.

However, Harding considers that social, ritual and political aspects of trading have been over-emphasised in ethnographic accounts of Melanesian societies and maintains that the primary function is economic. The Siassi, for example, "... engaged in social rituals of exchange as a means of acquiring valued need-serving goods" (1970:108). They acted as middlemen, exchanging goods at favourable rates by manipulating exchange ratios in the different Vitiaz Strait ports. A pig, for example, could be exchanged on Umboi for 5-10 packets of sago, which in turn were exchanged at Sio or Gitua for 50-100 pots. These pots could then be transported to New Britain, where they yielded 5-10 pigs (Harding 1970:139). Thus goods of little value in one community were transported to others where they were in short supply, or had high prestige, usually for ceremonial purposes, and thereby appeared to yield a considerable profit.

In both trade and exchanges of the more ceremonial variety, however, the concept of profit requires close examination, and Strathern (1983) points out some qualifications to the generally understood meaning of the term. In Melpa *moka* exchanges, for example, where a form of 'profit' appears to be built in to the system, material losses are nevertheless made up as gains in political prestige. Similarly, although surplus 'profits' of *tambu* shell may be acquired among the Tolai, they tend to be stored for subsequent redistribution. And in the *kula*,

The category of 'profit-making' tends ... to dissolve when considered more closely, into other categories. First, the 'profit' may have to be returned later. Second, it is in any case a by-product of competitive giving, the aim of which is to acquire renown. Third, it may simply represent the wish by the donors to maintain the partnership It is better, therefore, to refer to 'incremental giving' rather than to a 'principle of profit'.
(Strathern 1983:80)

In Morobe, too, the acquisition of profit does not seem to have been the fundamental aim. Harding notes (1970:105):

Ultimately the concept of profit appeared irrelevant because the objectives of trade for any participant community were highly specific. The Siassis counted the returns of trade in pigs and vegetable food which could be reserved for the 'feasts of merit' staged by local leaders. The Komba mountaineers were drawn to Sio because 'they smelled the grease of our coconuts', 'they were hungry for fish and salt'. Because of these specific objectives, transactions that might be judged unprofitable as measured by comparative inputs of labour time were desirable nonetheless.

In Siassi, any 'profit' gained was essential both for subsistence needs in their over-populated communities and for the prestige of men sponsoring feasts celebrating various stages of the life-cycle. But it was not merely an accumulation for individual use, and Freedman notes (1970:318):

The promotion of private accumulation, in Siassi, is ultimately in the public interest ... most private wealth is distributed in village-wide ceremonies.

Another aspect of trading in the Vitiaz network should be emphasised, and that is the role of trade partners. Although trading involved the manipulation of an awesome array of exchange ratios, it was not an impersonal affair designed to maximise acquisition at all cost, but involved a category of people who were regarded as kinsmen. Depending on relative age, trade partners would call one another 'brother', 'father' or 'son', and these relationships might be carried on in later generations until the distinction between 'real' and 'trade' kinship became blurred. Since transactions were carried out with a category of kin, the exact computation of exchange rates or enumeration of totals may not have been as important as it otherwise would have been. Nevertheless, it does seem that there was general agreement about exchange ratios; discrepancies recorded probably reflected the wide range in size and quality of the various commodities (Freedman

1970:154). Close trade partnership was associated not so much with differences in rates of exchange as with a time delay between initial gift and return (Freedman 1970:165).

The establishment of trade partnerships had a number of important advantages. In an area where warfare was endemic, a trade partnership afforded protection in an otherwise hostile community. Since exchanges were of such economic importance to both parties, this temporary suspension of hostility appears to have been strictly observed. The 'credit' extended in delayed return has been mentioned above. Pomponio (1983) stresses the role of trade partnerships in the extension of a kin network by prospective big-men in Mandok (Siassi). The establishment of a new trade partnership eventually transformed strangers into brothers, and an exchange of women in marriage commonly followed, thus creating affines. An additional link could be established by the adoption of children. What Pomponio refers to as "concepts of personal investments in people" (1983:181) were thus brought into operation as a means of establishing a group of people on whom a man could rely for support in various prestige-enhancing enterprises.

The importance of these trade partnerships is demonstrated by the strict etiquette observed by partners (Harding 1967:166). Hagglng over prices was not acceptable, and gifts had to be accepted graciously. Stealing of trade partners or undercutting of prices were likewise serious offences, while an underlying general principle of reciprocity avoided undue advantage being taken of the situation. The observance of this code of ethics suggests that the maintenance of amicable trading relationships was considerably more important than insisting on the most beneficial rates of exchange.

CONCLUSION

In the kind of cultural and economic climate outlined above it is suggested that an extensive abstract counting system was not an essential prerequisite for successful participation in most aspects of traditional life. Even in the relatively sophisticated trading network of the Vitiaz Strait, computation and enumeration appear to have been secondary considerations, while social factors such as the maintenance of harmonious trade relationships were essential for ensuring the supply of vital commodities and expanding kin networks in the incessant pursuit of prestige.

While the possession of a complete decimal series of numerals would certainly not be a disadvantage, it seems that the redundancy of such cultural items could have led to their gradual atrophy. It is thus postulated that the numeral series of Austronesian-speaking arrivals in north-east New Guinea gradually regressed under the influence of indigenous social and economic systems, and that their counting methods came to resemble those of their non-Austronesian-speaking neighbours.

REFERENCES

BIERSACK, Aleta

- 1982 The logic of misplaced concreteness: Paiela body-counting and the nature of the primitive mind. *American Anthropologist* 84/4:811-829.

DAHL, Otto Christian

- 1976 *Proto-Austronesian*. Scandinavian Institute of Asian Studies Monograph Series No.15. Lund: Curzon.

FREEDMAN, Michael P.

- 1967 The social and political organization of the Siassi Islands, New Guinea. Ph.D. thesis, University of Michigan.
- 1970 Social organization of a Siassi Island community. In Harding and Wallace, eds 1970:159-179.

HARDING, T.G.

- 1967 *Voyagers of the Vitiaz Strait*. Seattle: University of Washington Press.
- 1970 Trading in Northeast New Guinea. In Harding and Wallace, eds 1970: 94-111.

HARDING, T.G. and Ben J. WALLACE, eds

- 1970 *Cultures of the Pacific*. New York: Free Press.

LEACH, Jerry W. and E. LEACH, eds

- 1983 *The Kula: new perspectives on Massim exchange*. Cambridge University Press.

McELHANON, Kenneth A.

- 1984 *The languages of the Morobe Province: checklist, classification and field guide*. Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.

POMPONIO, Alice

- 1983 *Namor's Odyssey: education and development on Mandok Island, Papua New Guinea*. Ph.D. dissertation, Bryn Mawr. (Ann Arbor: University Microfilms International.)

STRATHERN, Andrew

- 1983 The Kula in comparative perspective. In Leach and Leach, eds 1983: 73-88.

THUNE, Carl E.

- 1978 Numbers and counting in Loboda: an example of a non-numerically oriented culture. *Papua New Guinea Journal of Education* 14 (special issue):69-80.

WURM, Stephen A. and B. WILSON

- 1975 *English finderlist of reconstructions in Austronesian languages (post-Brandstetter)*. PL, C-33.

POLICE MOTU AND THE SECOND WORLD WAR

Tom Dutton

1. INTRODUCTION¹

On 21 July 1942 Japanese military forces landed on the north-east coast of Papua² and immediately set out for Port Moresby overland across the Owen Stanley ranges. Soon afterwards other Japanese forces landed on islands further east and eventually on the mainland itself at Milne Bay³ - see map. In the ensuing attempt by allied forces to prevent the Japanese from extending their control to other areas of the country and to drive them from it large numbers of Papuans were drawn into the conflict.

During this time Police Motu, one of the two main lingue franche used in Papua before the war,⁴ was suddenly raised to the level of a language of strategic importance as it was used to maintain control over Papuans and to organise them to contribute to the allied war effort. As a result the language was described formally and recorded mechanically for the first time in its 50-odd year long history.

The present paper describes and discusses these wartime materials and uses them in association with other, more recently obtained, survey materials to gain insights into the nature of Police Motu before the war and hence to comment briefly on the more recent history of the language.

2. BACKGROUND: A BRIEF OUTLINE OF THE HISTORY OF POLICE MOTU

Police Motu is the name commonly used to refer to the language that is today officially known as Hiri Motu, the principal lingua franca of much of the former Territory of Papua, Papua New Guinea.⁵ It is a pidgin language based on Motu, the mother tongue of one of the two groups of original native inhabitants of the Port Moresby coastal area when Europeans first settled there, but now spoken by many others in the same area. Up until recently Police Motu was given no kind of official recognition or status although it served a major role in the development of the former British and Australian colony-cum-Territory as the unofficial language of administration. Indeed, for most of this time it was despised and discredited by expatriate authorities as a corrupt form of 'pure' or 'true' Motu. With the coming of independence to Papua New Guinea, however, its status improved so much so that it is now regarded as one of two unofficial national languages of the country.

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 351-406. *Pacific Linguistics*, C-94, 1986.

© Tom Dutton

with expanding government control over the country. From that time on the history of the language became one of increasing competition with the other main lingua franca spoken in Papua before the war, a "broken and rather rudimentary form of English".⁶ This English was a pidginised form of English that was the language of commerce and that has since been described as Papuan Pidgin English (Mühlhäusler 1978). It had been introduced into the country at its eastern and western ends by labourers recruited (sometimes illegally) to work in the marine industries in the Torres Straits and/or in their own areas and/or on sugar plantations in Queensland. It was also the first language of MacGregor's official police force and appears to have remained an important second language in the force for many years.⁷

Before the Second World War, then, Police Motu was spoken throughout most coastal areas of Papua and the immediate hinterland, the distribution reflecting time depth of contact and the pattern of economic development in an area. It was weakest in the eastern and western ends of the country where "English" was the language of commerce and where regional languages such as Suau in the eastern end and Kiwai in the western end were strong. It was also weak in the inland Central, Gulf and Western Divisions where government control was only progressively established after the war and where few, if any, able-bodied men had been away to work elsewhere in the country.⁸

Up to this time too the language had never been formally described or recorded although missionaries of the Seventh Day Adventist church used it for evangelical purposes in inland areas of the Central Division, and it was to be another 20 years before it was properly surveyed.⁹ Then it was surveyed by members of the Summer Institute of Linguistics for the government and the results were published (Brett et al 1962). In this survey the authors make two points which are of particular relevance to the forthcoming discussion. These are:

- (a) Although the surveyors found that there were "considerable local variations of Police Motu and numerous differences of pronunciation" (p.9) they concluded that there were only two dialects of Police Motu. These were a Central one, spoken in the Central District mainly by speakers of Austronesian languages that also happen to be closely related to Motu, and a non-Central one spoken elsewhere, but mainly by speakers of non-Austronesian languages, which, by definition are not related to Motu or any other Austronesian language in Papua or the Pacific. The non-Central dialect differs from the Central one in having a number of features that are not typical of that variety. For example, whereas the Central dialect generally followed the Motu manner of indicating possession with parts of the body and kinship terms (with suffixes such as *-gu my*, *-mu your*, *-na his, her, its*, etc.) as well as its manner of marking pronoun objects on the verb (with a set of suffixes similar to those used for marking possession (Lister-Turner and Clark n.d.a)) the non-Central dialect did not. Generally too the phonology of the Central dialect was closer to that of Motu. These two dialects will be symbolised CD and n-CD respectively and used hereafter as reference points in the discussion that follows;
- (b) "the type of Police Motu used in the Western District is intelligible to all Police Motu speakers" (p.11).

3. THE WARTIME MATERIALS

As already indicated Police Motu was only formally described and mechanically recorded for the first time in its life during the Second World War. There are thus two types of materials to be considered under this heading, written materials and oral materials. Each will be considered separately.

3.1 The written materials

When in 1941 war looked as if it might come to the south-west Pacific,¹⁰ and when in 1942 it actually did come, the Australian Army obtained copies of a number of "descriptions" of Police Motu for Army purposes. It is not known how many such descriptions were obtained or whether they were specifically commissioned, or not, by the Army, but three of them have been found in the archives of the Australian War Memorial in Canberra.¹¹ Two of these are anonymous (but presumably written by expatriate residents who knew the language)¹² and the third was by Maynard Lock, a pastor with the Seventh Day Adventist church in Papua.¹³

The three descriptions, excerpts of which are to be found herein in Appendices 1-3, are as follows:

- (a) *Handbook of Motuan (Police Motu) for use in Papua*. A printed booklet published by the Australian Army on 31 August 1941;
- (b) "Police Motu". An undated mimeograph by Maynard Lock;
- (c) "Police Motuan vocabulary". An undated mimeographed paper of unknown origin.

Hereafter these "descriptions" will be referred to as "The 1941 Handbook", "Lock's Police Motu" and "The Police Motuan Vocabulary". Each will be discussed in turn as each has its own characteristics and value before some general conclusions are drawn about the nature of the language they purport to describe. Before doing that, however, it is to be noted that:

- (a) the "descriptions" are barely more than wordlists containing little or no description or illustration of the phonology or grammar of the language or of variation in it;
- (b) none of the descriptions contains any illustrative text material given by a Papuan speaker.

For these reasons, as the language was generally despised by expatriates of the time (as overtly expressed by one of the authors of the descriptions to be considered) the descriptions, valuable as they are, are of lesser value than they might otherwise have been. They have, in particular, to be regarded as biased towards an expatriate overseer or "master" view of the language, and therefore do not accurately represent the language actually spoken by Papuans amongst themselves at the time, but only more or less so.

Taking each in turn and using the following abbreviations in addition to those already noted, for comparative purposes:

LTC	Lister-Turner and Clark (n.d.a)
PM	Police Motu
PoW-PM	Post-war Police Motu ¹⁴
M	('pure' or 'true') Motu

let us begin with the earliest one.

3.1.1 The 1941 Handbook

As already indicated this is a printed document that was published by the Australian Army before the war actually came to Papua. Its aim was "to present words and phrases in common use" (p.3) and it presents about 475 of these in an English-to-Police Motu listing. These words and phrases are spelled in an English-based orthography with the aim of indicating "as near as possible ... the pronunciation of the word" (p.3). Thus, for example, the Police Motu for *about (to go about)* is written 'giroa-low (ow as in plow)', *accept* as 'ubia (ub as in hub)', and *all* as 'ee-bone-ay' - see Appendix 1 - whereas in standard Police Motu orthography (which in turn is derived from Motu orthography (Lister-Turner and Clark, n.d.b)) these would have been written giroa lou, abia, and ibounai respectively. This manner of recording, though ambiguous in some cases, is nevertheless useful for present purposes for it shows how the variety of Police Motu being recorded was actually spoken and thus given an insight into its phonological nature.

The following are pertinent observations:

- (a) h is generally missing from those positions where in PoW-PM it is recorded:¹⁵

1941 PM	PoW-PM	Gloss
anaia	hanaia	<i>cross over</i>
dio	diho	<i>down</i>
isiu	hisiu	<i>star</i>
daka	dahaka	<i>what</i>
uiui	huihui	<i>hair</i>

or it is inserted where it is not recorded in PoW-PM:

hadavaia	adavaia	<i>to marry</i>
hatoa	atoa	<i>put</i>
haina	aena	<i>foot</i>

- (b) i is used for e:

aina	aena	<i>foot</i>
idau nigai	idau negai	<i>some time ago</i>
dabareri	dabarere	<i>daybreak</i>
masi	mase	<i>dead</i>

- (c) u is used for o:

tanu	tano	<i>ground</i>
lau	lao	<i>go</i>
gadubada	gadobada	<i>open sea</i>

- (d) w is used for v:

wadani	vadaeni	<i>all right</i>
wonagi	vanagi	<i>canoe</i>
weria	veria	<i>pull</i>

- (e) l is generally used for r:

gadala	gadara	<i>play</i>
ila	ira	<i>axe</i>
lakatania	rakatania	<i>leave</i>

1941 PM	PoW-PM	Gloss
lala	rara	<i>blood</i>
laronai	lalonai	<i>inside</i>
kailakaila	kahirakahira	<i>close</i>

(f) p is used for b and vice versa:

dipa	diba	<i>know</i>
kobukobu	kopukopu	<i>mud</i>
mata keburu	mata kepulu	<i>blind</i>

(g) k is used for g:

katoi	gatoi	<i>egg</i>
kunikai	gunika	<i>inland</i>
kuota	gwauta	<i>ten</i>
kwaremi	gwarume	<i>fish</i>
ia kwo	ia gwau	<i>he said</i>
maraki	maragi	<i>little</i>
kora	kwara	<i>head</i>

(h) some words have unexpected forms:

kaharaga	karaharaga	<i>quickly</i>
dogwata	dogoatao	<i>hold</i>
vadani	vadaeni	<i>all right</i>
hamaru	hamaoroa	<i>tell</i>
koa	koua	<i>shut</i>
kwaremi	gwarume	<i>fish</i>

In general the phonological characteristics of the vocabulary recorded in this volume are such that they suggest that the variety recorded was spoken by someone who was not a native speaker of Motu, and probably also not of any other Austronesian language. That is, it looks as if this material represents the speech of a non-Austronesian speaker.¹⁷

Grammatically very little is said in this Handbook about the structure of Police Motu and few sentences are given. From what is said about the use of personal suffixes on some nouns, however, it appears that the variety being described is that used in the Port Moresby area, for it is in this area that these suffixes have been used historically. Thus if this information is taken in association with that of the phonology just described it suggests that the variety being illustrated is some Port Moresby area non-Austronesian speaker's version of what is today called the Central dialect. Even so there are some interesting differences between this speech and that of the present-day speakers of this area. Thus for example:

(a) the focus markers *be* and *ese* of post-war Police Motu are not used where they are expected. Consider, for example:

1941 PM	PoW-PM	Gloss
inai daka?	inai be dahaka?	<i>What's this?</i>
dinagauna hida?	dinagauna be hida?	<i>What's the time?</i>
davana hida?	davana be hida?	<i>How much is it?</i>
inai laegu	inai be laegu	<i>This is mine.</i>
laegu X edisini?	laegu X be edeseni?	<i>Where is my X?</i>

- (b) the second person is not used in imperatives (though even in post-war Police Motu this depends to some extent on the situation). But consider:

1941 PM	PoW-PM	Gloss
oiemu ila mailaia!	oiemu ira oi mailaia!	<i>Bring your axe!</i>
naria namonamo!	oi naria namonamo!	<i>Wait patiently!</i>
dala hadibaia!	dala oi hadibaia!	<i>Point out the road!</i>
noho sisina!	oi noho sisina!	<i>Wait a bit!</i>

- (c) the anaphoric pronoun *ia* is not used where expected:

korana isiisi	(lauegu) kwarana ia hisihisi	<i>My head's aching.</i>
daika kwo?	daika ia gwau?	<i>Who said so?</i>

The vocabulary section of this Handbook and a comment on the vocabulary of the language in it are also interesting for the insight they throw on the lexical nature of the language of the time. According to the author (p.3) the vocabulary of Police Motu is to be regarded as identical to that of Motu except for "the importations of civilization" which "have no word in the Motuan tongue". In these cases the author advises, "English is adopted" and in that case the adoptions are given their "proper pronunciation".

In the vocabulary actually printed the following items occur which are different from present-day ones in some way, or are otherwise interesting. In this listing unless otherwise specified PoW-PM means 'in both dialects of PoW-PM' and all statements are more or less true as generalisations:

ITEM	GIVEN MEANING	COMMENT
anani	<i>fruit</i>	Probably a typographical error for <i>anina contents</i> . In PoW-PM <i>fruit</i> is <i>auhuahua(na)</i> .
aulolo	<i>mountain</i>	Counted as an idiosyncratic version of <i>ororo mountain, hill</i> in PoW-PM.
boidia	<i>to call someone</i>	A variant of <i>boiria to call someone</i> in PoW-PM.
botol	<i>bottle</i>	<i>bottle</i> is <i>botol</i> in PoW-PM, n-CD, <i>kavabu</i> in CD.
dekena	<i>on</i>	<i>on</i> is <i>dekena(i)</i> in PoW-PM, n-CD, <i>latanai</i> in CD.
dogwata	<i>hold</i>	<i>hold</i> is <i>dogwatau</i> (or something similar) in PoW-PM, n-CD, <i>dogoatao</i> in CD.
ena gena	<i>if</i>	<i>if</i> is <i>bema ... (negana(i))</i> in PoW-PM. This may be a transcription error for <i>inai nega this time</i> .
gabuli	<i>shark</i>	<i>shark</i> is <i>kwalaha</i> or <i>sak(i)</i> in PoW-PM, n-CD, <i>kwalaha</i> in CD. <i>gabuli</i> does not occur in Motu (LTC).
gadubada	<i>reef</i>	<i>reef</i> is <i>rip</i> in PoW-PM, n-CD, <i>moemoe</i> in CD. <i>gadobada</i> means <i>open sea</i> in PoW-PM.
hamaru	<i>report, tell</i>	PoW-PM, CD form is <i>hamaoroa</i> , n-CD <i>hamoroa</i> .
hanahanai	<i>always</i>	PoW-PM, CD form is <i>hanaihanai</i> , n-CD <i>hanahana(i)</i> .
idiena	<i>their</i>	PoW-PM, CD form is <i>idia-edia</i> , n-CD <i>idiena</i> .
ilama	<i>trochus shell</i>	<i>trochus shell</i> is <i>toea</i> or <i>trosel</i> in PoW-PM, n-CD, <i>toea</i> in CD. <i>ilama</i> is the Motu name for <i>trochus shell</i> before it is cut to make <i>toea armshells</i> .
itena	<i>our(incl.)</i>	PoW-PM, CD form is <i>ita-ena</i> or <i>isena</i> , n-CD <i>itena</i> .

ITEM	GIVEN MEANING	COMMENT
kapori	<i>turtle, tortoise</i>	PoW-PM form for <i>turtle</i> is matabudi. kapori does not occur in Motu (LTC) but means <i>Oh sorry</i> in Suau.
kavamasi kikaraia	<i>stupid, silly look up</i>	PoW-PM form is kavakava (momokani). PoW-PM form is koua (ki dekenai), lokaia or kia.
kiskis kamkam	<i>biscuit fowl</i>	PoW-PM form is biskesi, biskis or keke. PoW-PM form is kokoroku but kamkam also used in Milne Bay District for <i>bush fowl</i> and in other varieties for <i>fowl, bush fowl</i> for which the PoW-PM, CD word is kepoka.
kwaremi	<i>fish</i>	PoW-PM, CD form is gwarume, n-CD korume or korumi.
kworaia lo	<i>repeat</i>	PoW-PM, CD form is gwauraia lou, n-CD koraia lo.
levu	<i>skin disease</i>	In PoW-PM sipoma is the common form; levo means <i>tinea imbicata</i> or <i>ringworm</i> in Motu (LTC).
magi	<i>rubber</i>	PoW-PM form is raba. magi is the name for <i>Ficus Rigo</i> formerly used for obtaining wild rubber (LTC).
mahuta masi	<i>sleepy</i>	PoW-PM is matana ia metau or matana ia mahuta. In PoW-PM mahuta masi means <i>to sleep as if dead, to be dead to the world</i> .
namu noienoie otolo	<i>mosquito beg, cadge tree fern</i>	PoW-PM, CD form is namo, n-CD nemo. PoW-PM form is noinoi. No specific word for <i>tree fern</i> in PoW-PM. Presumably this is a variant of hatoro <i>cycas palm tree and fruit</i> in Motu (LTC).
otolu	<i>hill</i>	PoW-PM form is ororo. Not in Motu (LTC). May be transcription error for ororo.
pau	<i>bow (weapon)</i>	PoW-PM is peva. This is evidently a borrowing of English <i>bow</i> .
rata, susu	<i>milk</i>	PoW-PM, CD is rata, n-CD is susu. susu occurs in PoW-PM, CD in certain collocations such as magi/raba susu <i>latex</i> .
tanu wa siau	<i>ground boiling water</i>	PoW-PM form is tano. In PoW-PM wa siahu is <i>soup; boiling water</i> is ranu siahu.
wariwari	<i>mirror, looking glass</i>	PoW-PM form is (lukin)galasi. Motu for <i>mirror</i> is hevarihevari (LTC).

These items fall into the following categories:

(a) apparent transcription mistakes or variants of PoW-PM forms:

Apparent mistakes	Variants
anani <i>fruit</i>	boidia <i>to call someone</i>
aulolo <i>mountain</i>	kwaremi <i>fish</i>
noienoie <i>beg, cadge</i>	kwaraia lo <i>repeat</i>
pasi <i>sack, bag</i>	hamaru <i>report</i>
	hanahanai <i>always</i>
	itena <i>our</i>
	dekena <i>on</i>
	otolo <i>tree fern</i>

(b) forms that are not found in PoW-PM (and therefore may represent 'errors' or losses), or if found, have a different meaning from that indicated:

Not used	Different meaning
ena gena <i>if</i>	magi <i>rubber</i>
gabuli <i>shark</i>	kamkam <i>fowl</i>
kapori <i>turtle</i>	kapori <i>turtle</i>
gadubada <i>reef</i>	gadubada <i>reef</i>
ilama <i>trochus shell</i>	wa siau <i>boiling water</i>
kikaraia <i>lock up</i>	
kiskis <i>biscuit</i>	
kavamas <i>stupid, silly</i>	
otolu <i>hill</i>	
pau <i>bow (weapon)</i>	
levu <i>skin disease</i>	
wariwari <i>looking glass, mirror</i>	

(c) forms that are typical of one dialect of PoW-PM only:

n-CD	
botol	<i>bottle</i>
dekena	<i>on</i>
dogwata	<i>hold</i>
hanahanai	<i>always</i>
idiena	<i>their</i>
itena	<i>our</i>
susu	<i>milk</i>
kworaia lo	<i>repeat</i>
kwaremi	<i>fish</i>
hamaru	<i>report</i>

Given (i) that variants under (a) above are closer to n-CD forms than CD ones; (ii) that the occurrence of some of the items listed under 'not used' in (b) could be explained as having arisen because the informant or speaker from whom they were obtained was unfamiliar with Motu (e.g. such items as ena gena, gabuli, kapori, gadubada, otolu, pau) in which these items do not occur or occur with different meanings; (iii) that the forms under (c) are also n-CD ones; then it seems reasonable to assume that the vocabulary represents that of a non-Austronesian speaker trying to speak the Central dialect of Police Motu, an assumption which is consistent with that which was made independently on both phonological and grammatical grounds above. It is not possible to say, however, exactly where this speaker might have come from, but given that Police Motu was not widely known in the inland area west of Port Moresby he/she most likely came from the Koiari-Mountain Koiari area inland of Port Moresby or from neighbouring areas of Rigo to the east (Dutton 1969).

3.1.2 Lock's Police Motu

As already indicated this account was written by Pastor Lock of the Seventh Day Adventist church. He first came to live in Papua in 1924 and learned Police Motu and some Koiari at Bisiatubu on the Sogeri Plateau. Subsequently he went to work at Efogi in the mountains inland of Port Moresby where he learned Mountain Koiari, the vernacular of the people of that area and a close relative

of Koiari. During the war Pastor Lock was forced to leave Papua temporarily but returned as a member of the Australian army. Soon after his arrival his knowledge of the people and of their languages came to the notice of the authorities and he was taken into the Australian New Guinea Administrative Unit (or ANGAU) and eventually became responsible for broadcasts in Police Motu on the local radio station at Port Moresby. It was probably during this time that he wrote the account of Police Motu that is being discussed here. After the war Pastor Lock returned to his mission work and teaching until he was tragically killed during the sudden eruption of Mount Lamington in the Northern District in 1951.

Pastor Lock's account of Police Motu is interesting from a number of points of view. Chief amongst those is that he considered the language to be "not a real language" at all (p.3) but a "bastard language" (p.1) which had "no" (p.2) or "very few set rules" (p.1) - see Appendix 2. This was in some ways a surprising attitude to have expressed towards a language that his church used as a church language. It is not surprising, however, in terms of the negative attitudes held at that time towards pidgin languages in general and especially by officialdom in Papua.¹⁸ It is important, nevertheless, for present purposes in that it colours and/or determines the kind and amount of information that is included in this account about all aspects of the structure of the language. This is most serious in the phonology section where the description appears to have been taken directly from Lister-Turner and Clark's (n.d.a) published description of the grammar of that language. Consider, for example, the parallelism in the following statements:

LOCK, p.2

"The alphabet consists of nineteen letters: a, e, i, o, u, b, g, and ɣ, h, k, l, m, n, p, s, t, v, and two compound letters, kw and gw".

"The vowels have the Continental sound".

"The consonants are pronounced as in English but at times the ɣ has the ng sound".

"There are many Diphthongs, e.g.: ae, ai, ao, au, ei, eu, oi and ou".

"Not two consonants ever stand together".

"There is no rule for accent".

LISTER-TURNER AND CLARK, p.9

"Motu has been reduced to writing by means of an alphabet of 19 letters ... a, e, i, o, u, b, d, g, ɣ, h, k, l, m, n, p, r, s, t, v. There are also two compound letters: kw and gw".

"The VOWELS have the "continental" sounds".

"The CONSONANTS are pronounced as in English, except for ɣ...".

"There are many DIPHTHONGS, e.g. ae, ai, ao, au, ei, eu, oe, oi, ou".

"No two consonants ever stand together".

"There is no fixed rule for Accent or Stress".

As a result this account contains no information about phonological aspects of the language, that is, about how words were actually pronounced at the time, because the standardised spelling of the words masks differences in pronunciation. Consequently it contains no information about variation or dialects. In addition the nature of the account means that the information in it has to be treated with a certain amount of suspicion, a suspicion strengthened by the fact that the vocabulary also contains a large number of words taken directly from Motu but which were not commonly used in post-war Police Motu and still

are not, except by those who knew, or know, Motu, or lived, or live, in close contact with the Motu. Thus, included in this vocabulary are items such as *ariara natuna illegitimate child*, *ariara hahine prostitute*, *ahetoni parting with friend*, *au burena flower*, *ava raua to weed*, *baragi lungs*, *be why?*, *biru gardening*, *daiguna corner*, *darere overcome*, *weak*, *edeheto how?*, Motu numerals (at least above *ten*) and colour terms (p.32) - see Appendix 2.

Even so, by comparing the contents of this vocabulary with Motu and other later forms of Police Motu and by taking into account other information obtained from Pastor Lock's surviving brother - see note 13 - it is possible to separate out those aspects of the language illustrated here that were probably not, and those that probably were, genuinely part of the variety being described here at the time. Thus the following can probably be accepted as being part of the grammar of the variety being illustrated:

- (a) nouns were invariable for number except for a very few which had the first syllable duplicated in the plural, e.g. *mero boy* becomes *memero boys*, *tau man* becomes *tatau men*;
- (b) *be* was used as a focus marker as in PoW-PM, e.g. *unai be daika ena natuna? Whose child is that?* (p.2), *hari dina be edeseni oi lao? Where are you going today?* (p.3), *lauegu tamana be edeseni? Where's my father?* (p.3);
- (c) *vadaeni* was used as a recapitulator in connected discourse or complex sentence as in PoW-PM, e.g.:
- ai lao unai motumotu dekenai, vadaeni kerukeru ai giroa
we(excl.) go that island to OK tomorrow we(excl.) turn.around
 lou
again
We are going to that island and will return tomorrow (p.3);
- (d) *vadaeni* was also used as a completive aspect marker as in PoW-PM, e.g.:
- ruma haginia tauna ia heau vadaeni
house erect man he run finish
The man who built the house has run away (p.3);
- (e) relative clauses were formed on a head noun (e.g. *tauna man, person*) as in PoW-PM:
- ruma haginia tauna ia heau vadaeni
house erect man he run finish
The man who built the house has run away (p.3);
- motuka dekenai ia moru tauna, harihari sisina ia namo
car from he fell man now little.bit he good
The man who fell off the truck is now getting better (p.3);
- (f) there was a distinction between inclusive and exclusive forms of the first person plural pronoun, viz. *ita we(incl.)* and *ai we(excl.)*;
- (g) *dohore* was used as a future tense marker as in PoW-PM:
- dohore hari dina lau lao digudigu
future this day I go bathing
I'm going for a swim today (p.3);
- (h) the verb *to have* was expressed by pronoun + *dekenai* + *noho* as in PoW-PM:

- oi dekenai aniani haida ia noho
you at food some it stays
Have you got anything to eat? (p.3);
- (i) yes-no questions were formed by phonological means and not by changes in word order or morphological means as in PoW-PM:
- oi gorere?
you sick
Are you sick? (p.31);
- (j) there was no agreement between subject and verb as in PoW-PM:
- lau mai
I come
I am coming, I come
- oi mai
you come
you are coming, you come
- ia mai
he come
he is coming, he comes (p.31);
- (k) the one verb form was used for simple present, present continuous and past tenses as in PoW-PM. Thus lau mai meant *I come, I am coming, I came (p.31);*
- (l) want or volition were expressed by ura as in PoW-PM:
- lau ura oi lao bona oi tahua
I want you go and you search
I want you to go and look for it (p.31);
- (m) bona was used for the conjunction *and* between clauses as in PoW-PM. Consider the example in (l) above and the following:
- oi lao bona oi karaia!
you go and you do.it
You go and do it! (p.31);
- (n) baine ... neganai was used for *if* (which is not the same as in PoW-PM):
- baine lau hamaoromu neganai ...
if I tell.you time
If I tell you (to do something), you ... (p.31);
- (o) the second person pronouns oi *you*(sg.) and umui *you*(pl.) were used in imperatives as in PoW-PM:
- oi lao bona oi karaia!
you go and you do.it
You(sg.) go and do it!
- umui helai diho!
you(pl.) sit down
You(pl.) sit down!
- (p) suffixes on verbs were used to indicate pronoun objects as in the Central dialect of PoW-PM:
- baine lau hamaoromu neganai ...
if I tell.you time
If I tell you (to do something), you ...

- (q) *ia*, the third person singular pronoun, was used as an anaphoric subject pronoun when the subject was a noun phrase:

unai mero maragi laegu tadina ia botaia
that boy small my brother he hit
That little boy hit my young brother (p.2);

laegu kakana be kahirakahira ia mase
my brother focus close he die
My brother is dying (i.e. almost dead) (p.3).

These features are consistent with both dialects of Police Motu as defined in 1961 except for features (n) and (p). Because this latter is the chief defining feature of the Central Dialect of the language the variety described by Lock most probably represents that dialect also.

Before leaving this account, however, there is one other aspect of it that needs mentioning. That is that the particular attitudes expressed by the author lead him to a number of general observations about the nature of the language that are important for present purposes. These observations are:

- (a) the language takes its name from the Native Police Force whose members "picked up more Motu words and added them to the Bastard language" (p.1);
- (b) Police Motu is difficult for many 'pure' Motu speakers to understand unless they have had much practice. "It sounds as bad to them as "Pigin" [sic] English does to us and it is really a bastard language" (p.1);
- (c) "Nearly all the different tribes have their own formation of Police Motu sentences" (p.3).

Although the significance of these observations is not immediately obvious it becomes so once other information to be presented shortly is taken into account. We return to this in the concluding section of this paper.

3.1.3 The "Police Motuan Vocabulary"

This "Vocabulary" consists of two parts: (a) a vocabulary consisting of an English-Police Motu list of about 300 items and a Police Motu-English one of about 650 items, and (b) a set of 84 sentences divided up into groups under the following headings: "Of Building (15), Of the Sea (9), Medical (10), General (16), The Track-Carrying (18), and Work Generally (16)" - see Appendix 3. There is, however, no introduction to this volume such as is to be found in the 1941 Handbook or Lock's "Police Motu", so that there are no comments in it about attitudes to it or about its distribution, general nature, phonology, etc. The work is similar to Lock's just reviewed, however, in that:

- (a) it generally uses Motu spellings for words and grammatical elements;
- (b) it is Motu-oriented in vocabulary (although not as strongly as Lock's). Thus, for example, it contains such Motu forms as *ariara hahinena prostitute*, *auna timber*, *lesi drain*, *petaia rock back and forth*, *au tahona a splinter*, *au todena sap or gum* that have never been used by most Police Motu speakers.

It differs from Lock's, however, and the 1941 Handbook in:

- (a) having a more elaborate Police Motu-English vocabulary;

- (b) that it does not use the ɣ (representing a voiced velar fricative);
- (c) that it is not always consistent, e.g. *leg* is given as *aina* instead of the expected (in the author's terms) *aena*, *gwada* for *strong* instead of the expected *goada*;
- (d) providing a much more elaborate set of sentences in the language - indeed the best so far available from this period - which agree in structure with those in Lock's account (where the two overlap, that is), except in the following respects:
- (i) *bema* is used for *if*:
- bema au dorina ia mareremarere vadaeni ia dika*
if tree top/end it shakes/moves OK it no.good
If the post moves when pushed it is no good.
- (ii) the focus markers *be* and *ese* are not used;
- (iii) the second person pronoun is not used in imperatives.

Significantly this account agrees with the 1941 Handbook in two of these features (notably (ii) and (iii)) and therefore raises questions about the occurrence of these features in different varieties of Police Motu at the time. However, since these features are part of every variety of post-war Police Motu it is assumed that their non-occurrence in this material is to be explained on other grounds. Otherwise the sentences presented in this account show that the following features were part of the grammar of the variety of Police Motu being illustrated in addition to those manifested in Lock's account:

- (a) adjectives have suffixes *-na* and *-dia* attached to them to indicate singular and plural number of the noun head, e.g.:
- au latadia*
tree long.pl.
long logs ,
- kurukuru kwadogina*
grass short.sg.
short grass
- (b) *-ai* is used to indicate direction towards or location: *udai* *into the bush*, *kone-ai* *on the beach*, *vanagi-ai* *on the canoe*;
- (c) negative imperatives are formed by placing *lasi* after the verb, e.g.:
- kurukuru kwadogina utua lasi!*
grass short.sg. cut not
Don't cut the short grass!
- (d) *sibona* is used for *only*:
- kurukuru latana sibona*
grass long.sg. only
only long grass
- (e) adverbs come after the verb:
- kwatua namunamu*
tie securely
tie securely

geia dobudobu
dig deeply
dig deeply

- (f) garina is used for *lest*:

medu ia diho garina
rain it down lest
lest the rain come down

- (g) a is used as a yes-no question tag:

oi gorere a?
you sick eh
Are you sick? (You're sick aren't you?)

- (h) edeheto is used for *how* in some sentences:

oi edeheto a?
you how eh
What is the matter with you?

- (i) gwauraia is used as an intentive aspect marker:

idia huaia gwauraia
they carry about.to
They're preparing to/about to carry

From what has been said it would appear that the variety described in this account is again a variety of the Central dialect. There are two small slips of the pen, however, which suggest that this impression is artificially created by using Motu spelling and vocabulary. These two slips are *umi you(pl.)* and *kwarume fish*. The former is the distinctive form of the Motu second person plural pronoun *umui* used in the non-Central dialect today and *kwarume* is more like a non-Central form of Motu *kwarume fish* than a Central dialect one - compare the 1941 Handbook form *kwarem* *fish* already referred to. As a result it raises suspicions that the account is really based on some non-Austronesian language speaker's version of the Central dialect (principally because object suffixes are used on verbs). If so this would make the variety something similar to that described in the 1941 Handbook and probably also that which underlies Lock's account given that he worked mainly in a non-Austronesian language speaking area.

3.1.4 Some concluding remarks

Now that the three wartime "descriptions" have been presented and discussed the question arises: What do they tell us about the nature of Police Motu at the outbreak of war, and by implication immediately before? Clearly a great deal, although perhaps not as much as they could have had they not varied so much in quality and quantity of data presented. Most importantly, however, the three are in reasonable agreement with one another in most respects, once allowance is made for a tendency to ('pure') Motu-ise everything described, and appear to describe a variety that was most probably spoken by non-Austronesian language speakers in the Port Moresby area, that is, a non-Austronesian speaker's version of the Central dialect. This variety is recognisably the same as that still spoken in Papua New Guinea today. Where the "descriptions" differ from one another, however, and/or raise questions about the nature of Police Motu at the time (because they were written by untrained linguists and do not contain material actually recorded by Papuan speakers) is in the following respects:

- (a) in respect of the position of the following grammatical features:
- (i) the use of dekena versus -ai as direction and/or location markers;
 - (ii) the use of bema versus baine ... neganai for *if*;
 - (iii) the use of be and ese as focus markers;
 - (iv) the use of second person pronouns in imperatives;
 - (v) the use of ia as an anaphoric pronoun subject in sentences;
 - (vi) the use of edeheto *how*.
- (b) in respect of its actual vocabulary content. More specifically what is the status of the 1941 Handbook's "odd" vocabulary that was listed and discussed above? What was the status of the Motu vocabulary that appears in Lock's "Police Motu" and the "Police Motuan Vocabulary" that is not normally used in post-war Police Motu?
- (c) in respect of its variation between and within different areas.

We return to these questions once the wartime oral material has been presented in the next section.

3.2 The oral material

3.2.1 Introduction

When war came to Papua in 1942 - it began with the bombing of Port Moresby in February 1942 - there was general panic (Nelson 1982) and a real danger that the Allied Forces would lose control of the Papuan population with disastrous consequences for any planned attempt to halt the southward movement of the Japanese Forces. Consequently, as part of their effort to keep Papuans calm and under control and on the side of the allies, a small party of them was taken to Australia to see the allied war effort. In particular they were taken there to give them the impression that Australia was "one vast armed camp" and that "the Allied arms must in the long run prevail"¹⁹ (A.W.M. 431/8/3:2). The members of the party were then expected to return to their own areas to spread the good news about what they had seen in Australia. They were also expected to make gramophone recordings for the same purpose in Police Motu and other languages for distribution to ships' captains, villagers, and others who had gramophones, including those behind enemy lines.

The party finally selected consisted of six Papuan constables (one each from the Gulf and North-Eastern Divisions and two each from the Delta and Central Divisions) and a sergeant "from T.N.G. [Territory of New Guinea] Police"²⁰ (A.W.M. 431/8/3:2). They were gathered together briefly in Port Moresby and then trans-shipped to Townsville, in North Queensland.²¹ From there they travelled to Brisbane by train and eventually to Sydney and beyond by train and plane. In the process the party was shown army camps, ordinance stores, tanks and other weapons in use, airfields, farms, factories of different kinds - see Plate 3 - tall buildings with automatic lifts and other city wonders. They fired different kinds of guns and went for a ride in a tank and in an aeroplane - see Plate 4 - and as a grand finale were introduced to General MacArthur in Sydney. Subsequently they were taken to the Australian Broadcasting Commission studios in Melbourne where some of them made recordings in Police Motu and other languages. The Police Motu recording was made by Nanai Gigovi from Babaguna village near Kikori in the Gulf of Papua - see Plates 1-4. A copy of this recording is to be found in the archives of the Australian War Memorial, Canberra.

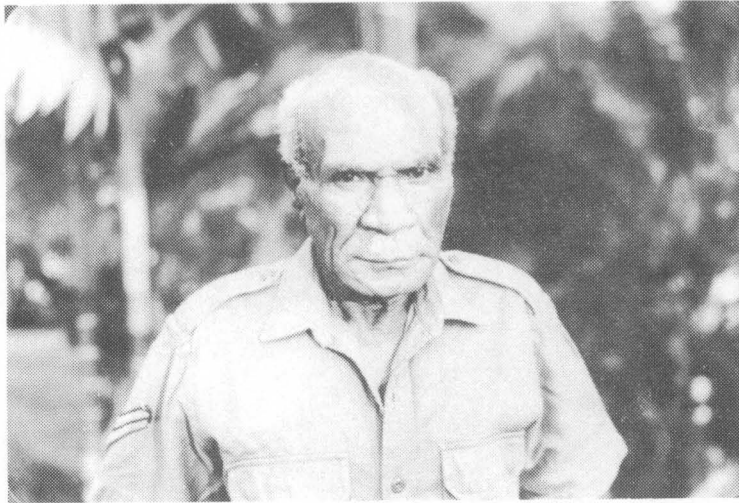


Plate 1: Nanai Gigovi, of Babaguna village, showing stripes obtained during World War II. (The shirt, however, is of the present day.)



Plate 2: Nanai Gigovi in June 1984



Plate 3: In a munitions factory, September 1942 (left to right: Nanai Egi (Tubuseleia village), Dedewa (Dorom, Rigo), Bauba Maba (Mambare R.?), De'ari Maba (Gulf), Nanai Gigovi (Kikori), Sgt.Mj.Pokonau (Manus Is.) [Missing sick was Kivovia from Vailala, Gulf]



Plate 4: Going for a flight (left to right: Mr Cowley (RAAF - killed in Papua), Bauba (Mambare), Nanai Egi (Tubuseleia), Capt.Stern/Still(?), Nanai Gigovi (Kikori), De'ari Maba (Gulf), Dedewa (Rigo), Mr Stanley (Geologist from Papua).) [Missing sick was Kivovia from Opa village, Vailala R. area, Gulf]

3.2.2 Nanai Gigovi's text

The recording made by Nanai Gigovi is 30 minutes long and consists of a number of tracks in Police Motu and one in Nanai's mother tongue Kerewo. The Police Motu tracks last for 25 minutes although because part of one track is repeated this section is only effectively 20 minutes long. The Kerewo track runs for five minutes and is a parallel one to the first part of the Police Motu recording.

In the recording Nanai is interviewed by another speaker of Police Motu whose identity is not known but who appears to be a native speaker of English. The interviewer's questions are short and do not occupy more than about two minutes of the recording time. Nanai's text was unscripted and he spoke freely and off the cuff. This recording has been transcribed and checked with Nanai.²² An extract from this transcript is included herein as Appendix 4.

3.2.3 Linguistic features of Nanai's text

Nanai's speech is characterised by the following linguistic features:

(a) in phonology:

- (i) he does not use the full range of Motu sounds. In particular he does not use *g* or either of the diphthongs *ao* and *ae*, substituting *g* or *k* for the former, and *au* and *ai* for the latter respectively. Nor does he pronounce *gwa* as *gwa* but rather as *kwo* or *ko* (e.g. *koraia* or *kworaia* for *gwauraia*);
- (ii) he has distinctive pronunciations for some words, e.g. he says *hanahanai* for *hanaihanai*, *hekula* for *hekure*, *idiena* or *idia-edia-ena* for *idia-edia*, *gaukwoaraia* for *gaukaraia*, and *omi* for *umui*;
- (iii) he substitutes *l* for *r* and vice versa, and *w* for *v*;
- (iv) he often 'drops' *h* as in *dio* for *diho*, *lai* for *lahi*, *daka* for *dahaka*;
- (v) he uses a rhythm and intonation pattern more characteristic of his own language, Kerewo, than of Motu.²³

(b) in grammar he uses the same structures as have been described in the wartime written materials above where these overlap with one another. Where they do not overlap (as has been indicated in section 3.1.4 above) he:

- (i) uses *dekena* as a general direction-cum-location marker and never *-ai*, e.g. *kone dekena on the coast*;
- (ii) uses *bema ... (negana)* for *if* and not *baine ... neganai*, e.g.

bema lau noho lagani tamona negana gabu ibounai lau itaia ia ore
if I stay year one time place all I see it finish
If I'd have stayed for a year I would have (completely) seen every-
thing.
- (iii) uses *be* and *ese* as focus markers, e.g.

idia gwau ... hari be mase taudia idia karaia
they say now focus dead people they make
They said ... that it was our ancestors who were making the things
now.

Diapan ese idia stat
Japan focus they start
The Japanese started it

- (iv) uses subject pronouns in imperatives, e.g.

Umui naria! Dohore ai mai.
you(pl.) wait future we come
You wait! We'll come.

Dohore! Oi naria!
later you wait
Later on! You wait (and see)!

- (v) sometimes uses the plural object suffix *-dia*²⁴ on verbs, e.g.
 compare:

diba taudia ia siaia
know people they send
they sent people who know

lau davaradia
I found.them
I found them.

Besides the structures he has in common with the written materials the following extra ones are exemplified in the text:

- (i) *ore* or *ia ore* following verbs as a completive marker indicating that everything has been done or seen, etc.

idia diba ore, gau ibounai
they know finish thing all
They know all about everything

gabu ibounai lau itaia ia ore
place all I see it finish
I would have seen everything completely

- (ii) *traim* is used as a verb in its own right and in combination with other verbs to indicate *to try to do something*, e.g.

lau ura traim
I want try
I wanted to try it.

ai faia ai traim traim
we fire we try try
We tried and tried to fire

- (iii) *ura* is used before other verbs to indicate volition or desire, e.g.

lau ura hereva
I want say
I want to say

lau ura traim
I want try
I wanted to try it

- (iv) *badina* is used as a subordinating conjunction for *because*:

badina idia laloa Pafuom be idia be goada lasi
because they think Papuans focus they focus strong not
because they think that Papuans are not strong

- (v) the third person singular pronoun *ia* is often used for the corresponding plural one *idia* as an anaphoric subject pronoun before verbs, e.g.

diba taudia ia siaia
know people they send
They sent people who know

- (c) in vocabulary he uses a number of English derived forms, some of which are distinctly Pidgin English in character. The full list of these forms is (in order of occurrence in the text):

ITEM	MEANING	CONTEXT
<i>misinigani</i>	<i>machine gun</i>	<i>misinigani idia karaia</i> <i>machine.gun they make.them</i> <i>they make machine guns</i>
<i>wosipi</i>	<i>warship</i>	<i>wosipi badadia idia karaia</i> <i>warship big.ones they make.them</i> <i>they make big warships</i>
<i>plaimasisi,</i> <i>plaimasini</i>	<i>aeroplane</i>	<i>plaimasisi ibounai idia karaia</i> <i>aeroplane all they make.them</i> <i>they make all the aeroplanes</i>
<i>mistim</i>	<i>mix</i>	<i>tano bona auri idia nadua idia mistim</i> <i>soil and iron they cook they mix-them</i> <i>they cook up soil and iron and mix them</i> <i>together</i>
<i>laidi</i>	<i>light</i>	<i>ia daekau inai laidi bamona</i> <i>it goes.up this light like</i> <i>it goes up like this light</i>
<i>tanika</i>	<i>tank</i>	<i>idia bubua tanika dekena</i> <i>they tip.it tank into</i> <i>they tip it into the tank</i>
<i>misin</i>	<i>machine</i>	<i>idia abia lao misin dekena</i> <i>they get.it go machine in</i> <i>they take it in a machine</i>
<i>gavamani</i>	<i>government</i>	<i>gavamani ia bogahisi maraki lasi</i> <i>government it sad little not</i> <i>the government was very sad</i>
<i>Ostrelya</i>	<i>Australia</i>	<i>Ostrelya taudia idia bogahisi maraki las</i> <i>Australia people they sad little not</i> <i>the Australian people were very sad</i>
<i>Pafuom</i>	<i>Papuans</i>	<i>idia lalao Pafuom be idia be</i> <i>they think Papuans focus they focus</i> <i>goada lasi</i> <i>strong not</i> <i>they thought that Papuans were not strong</i>
<i>Gavana</i>	<i>governor</i>	<i>Gavana ia bogahisi</i> <i>governor he sad</i> <i>the Governor was sad</i>

ITEM	MEANING	CONTEXT
plaua	<i>bread</i>	idia karaia plaua they make bread they make bread
raisi	<i>rice</i>	raisi idia karaia rice they make.it they make rice
bolomakau	<i>cattle</i>	bolomakau momo cattle plenty there were plenty of cattle
sipi	<i>sheep</i>	sipi momo sheep plenty there were plenty of sheep
nambawan	<i>first</i>	nambawan be Taunsvulo first focus Townsville the first one is Townsville
Taunsvulo	<i>Townsville</i>	nambawan be Taunsvulo first focus Townsville the first one is Townsville
seken	<i>second</i>	seken be Bulespen second focus Brisbane the second is Brisbane
Bulespen	<i>Brisbane</i>	seken be Bulespen second focus Brisbane the second is Brisbane
Sidin	<i>Sydney</i>	ia mai Sidin it come Sydney then there's Sydney
tauni	<i>town</i>	tauni kavakava town mad there are towns all over the place
faketri	<i>factory</i>	faketri momo factory plenty there were plenty of factories
wailisi	<i>wireless,radio</i>	wailisi dekena idia sikulu wireless about they learn they learn about wirelesses/radio
sikulu	<i>learn</i>	wailisi dekena idia sikulu wireless about they learn they learn about wirelesses/radio
stat	<i>begin,start</i>	Japan taudia ia stat Japan people they start the Japanese started it
Japan,Diapan	<i>Japan</i>	Japan taudia ia stat Japan people they start the Japanese started it

ITEM	MEANING	CONTEXT
tren,teleni	<i>train</i>	tren dekenai lau lao::: train in I go: : : I went a long way in the train
kempa	<i>camp</i>	inai kempa ibounai inai tau kurokuro this camp all this man white kavakava mad in all these camps there were white men all over the place
ami	<i>army</i>	ami solodia army soldier army soldier
solodia	<i>soldier</i>	solodia momo herea soldier plenty very there were very many soldiers
aut	<i>out, finish, leave</i>	memero matamata ia sikulu dekena ia boys new they school at they aut, ia gaukaraia finish they manufacture when new boys leave school they become producers
bomu	<i>bomb</i>	bomu ia negea Hanuabada dekena bomb they throw Hanuabada at they dropped bombs on Hanuabada
stanibai	<i>be prepared, ready</i>	ai danu ai stanibai we also we prepared/ready we were also prepared
Inglisi	<i>English</i>	lau lalao be Inglisi be momo herea I think focus English focus plenty very I think there are many more English (than Japanese)
hapu	<i>half, part</i>	hapu be idia negea half focus they throw.away they throw away half/part of it
elopen	<i>aeroplane</i>	elopen danu idia karaia aeroplane also they make.them they also make aeroplanes
injini	<i>engine</i>	injini lasi engine not they don't have any engine (in them)
bensini	<i>petrol</i>	O, bensini be halusia oh petrol focus cause.lose Oh, they waste petrol
halusia	<i>waste</i>	O, bensini be halusia oh petrol focus cause.lose Oh, they waste petrol

ITEM	MEANING	CONTEXT
motuka	<i>motor car</i>	ia motuka ese ia be [cut off] they car subject they focus they ... motor car ...
miti	<i>meat</i>	miti momo meat plenty there's plenty of meat
tin(i)	<i>tin</i>	tini dekena tin in in the tin
Amerika	<i>American</i>	Amerika danu lau dav - Amerika kini lau America also I find America king I lao itaia. go see.him I came upon Americans also - I went and saw the American commander
poren	<i>parade, line up, fall in on parade</i>	memero ibounai ai poren boys all we fall.in all of us boys paraded
Papua	<i>Papua</i>	ai be Papua we focus Papua we are Papuans
faia	<i>shoot, fire</i>	pidi danu lau faia rifle also I fired I also fired a rifle
blenigani	<i>brengun</i>	blenigani lau faia brengun I fired I fired a brengun
traim	<i>try</i>	ai faia ai traim traim we fire we try try we tried and tried to fire them
	<i>test</i>	lau umui traim I you(pl.) test I'm testing you(pl.)
tanika	<i>tank (of war)</i>	lau lao tanika gabuna dekena I go tank place to I went to where the tanks were
et thausen	<i>eight thousand</i>	et thausen idia gaukara iniseni eight thousand they work here eight thousand were working there
waia	<i>wire</i>	waia danu idia karaia wire also they make.it they also make wire
babuwaia	<i>barbed wire</i>	babuwaia danu idia karaia barbed wire also they make.it they also make barbed wire

Of these about two thirds are names for items that are foreign to Papuan culture and which were introduced by the intrusive colonial culture at different times. Thus there are names of towns and countries (Ostrelya, Pafuom, Taunsvulo, Bulespen, Sidin, Inglisi, Amerika, Papua), various kinds of machines and weapons (injini, motoka, tren, misin, fakteri, wailisi, plaimasini/eloplen, masinigani, blengani, wosipi, bomu(a), tanika), foodstuffs (plaua, raisi, bolomakau, sipi) and other miscellaneous items (tanika, tin, waia, babuwaia, Gavmani, Gavana, ami, solodia). All of these are clearly borrowings from English of one kind or another and can be dated fairly accurately according to specific cultural events, e.g. some such as Ostrelya, Sidin, Bulespen, Gavamani, Gavana, plaua, raisi, bolomakau, sip, tin, waia probably can be assigned unambiguously to the earliest phase of contact (although plaua, bolomakau and sipi are unexpected forms and will be returned to later). Others such as tauni, Papua, fakteri, injini, misini, tanika, waia and babuwaia have most probably to be assigned to a later period (e.g. early 1900s) when industrialisation began to be felt in Papua as mining and the plantation industry began to be developed and when British New Guinea was renamed Papua. Others again probably belong to later periods, e.g. 1920-30s for motoka, bensini, plaimisini/eloplen, and 1942 for the war related items of masinigani, blenigani, tanika, bomu(a), wailisi, ami, solodia, Amerika, Japan.

As indicated these items are straight borrowings from English and as such are generally uninteresting because of their form and because they are so expected. Yet it is important to recognise that all except the war related ones must be regarded as thoroughly integrated into Police Motu before the war, that is, they had been in use so long that they must be regarded as part of the language and no longer as borrowings sitting uncomfortably in the language. We return to this again later. Meanwhile there are a few items of this set which, because of their form, are interesting in other ways and are in need of some further discussion. These items are tauni, masinigani, blenigani, wosipi, motoka, plaimasini/eloplen, plaua, bolomakau, sipi.

Of these tauni, masinigani, blenigani and sipi are interesting because there were other words in Police Motu at the time that could have been used for these items. Thus, for example:

- (i) the Motu word *hanua* *village* covers the town concept and Nanai uses it in fact as such in the recorded text, e.g.

kone	dekena	hanua	ia	noho	...	hanua	badaherea	...	hanua	kavakava
<i>beach at</i>	<i>town</i>	<i>they stay</i>		<i>town</i>	<i>big.very</i>		<i>town</i>	<i>mad</i>		
<i>there are towns on the coast</i>	...	<i>big towns</i>	...	<i>towns</i>	<i>everywhere</i>					
- (ii) the Police Motu word *ipidi* or *pidigauna* *gun* could have been used in combination with descriptive adjectives or phrases to give periphrastic equivalents of masinigani and blenigani;
- (iii) the Police Motu word *mamoe* (itself a borrowing of a common mission term for *sheep* in Polynesia) could have been used for sipi *sheep*. In fact Nanai uses both in this recording.

Of the remainder bolomakau is interesting because it is a borrowing of a common Pidgin English form for *beef*, *cattle*, plaua because it is not beredi (or something similar), wosipi, motoka and eloplen because they are not manawo, lori and plaimasini respectively.

But to return to the remaining vocabulary in the listing given above, that is, the third that are not names for items foreign to Papuan culture. These are (following the same order as in the listing):

NANAI'S PM	MEANING	MOTU EQUIVALENT (LTC)	COMMON OR POSSIBLE PoW-PM FORMS
mistim	<i>mix</i>	buloa	giroagiroya
laidi	<i>light</i>	diari	diari, lahi
nambawan	<i>first</i>	ginigunana	ginigunana
seken	<i>second</i>	iharuana	iharuana
sikulu	<i>learn, school</i>	hadibaia	hadibaia
stat	<i>begin, start</i>	hamatamaia	hamatamaia
kempa	<i>camp</i>	taruka	hanua maraki maraki ta, ruma ta
stanibai	<i>be prepared, ready</i>	hagoeva-ia, nari-a	hagoevaia, naria
traim	<i>try, attempt</i>	dibaia (<i>to try a thing</i>) karaiatoho (<i>to try to do</i>)	verb + itaia
hap(u)	<i>half, part</i>	kaha-na	kahana, duduna
halusia	<i>waste</i>	petapetalaia	negea kava, inua kava
miti	<i>meat</i>	vamu	vamu
poren	<i>parade, line up, fall in on parade</i>	? (varo karaia(?))	? (varo karaia(?))
faia	<i>shoot, fire</i>	pidia	pidia
et thausen	<i>eight thousand</i>	daha-taurahani	momo, momo herea, hutuma bada herea

In all except poren *to parade* there is a perfectly good Motu word which could have been borrowed by Police Motu speakers if they had consciously sought to expand their vocabulary using Motu as a base (as expected in the Central dialect and as implied in the definition of Police Motu as a simplified form of Motu), or a perfectly good Police Motu form or expression which could have been invented if the resources of the basic vocabulary of Police Motu were to be used to the full (as is implied in the definition of pidgin languages as languages making maximum use of minimum means). This can be seen from the above chart in the right-hand column. What then is the explanation for the occurrence of these words, some of which (e.g. mistim, nambawan, stanibai, traime) have a distinctly Pidgin English form?

There are two possible explanations:

- (a) they were not part of Police Motu at the time but were restricted to Nanai Gigovi's speech for some reason;
- (b) they were part of Police Motu at the time and not peculiar to Nanai Gigovi's speech.

There is no way of deciding between these two possibilities on purely formal grounds. The only way that this can be done is by (a) looking more closely at Nanai's background to see if there is some explanation to be found there for the use of these forms and/or (b) trying to determine by survey means whether these forms are also used by other Police Motu speakers of the same generation.

3.2.4 The source of Nanai's English-based vocabulary

3.2.4.1 Nanai's background

As already indicated Nanai comes from Babaguna village near Kikori in the Gulf of Papua. He grew up speaking Kerewo, a Kiwaian language (Wurm 1975:327-328) and when he was old enough went to work for Europeans in various capacities ranging from houseboy-cum-cook to labourer, launch driver and government interpreter. During this period he learned to speak 'English' and Police Motu. When war came to Papua he was engaged to help recruit labour from the Kikori area for the war effort before being summoned to Port Moresby to go on the propaganda tour already described. After he returned from Australia he served in the Papuan Infantry Battalion, attaining the rank of corporal. Following the cessation of hostilities he returned to his home area where he has remained apart from further short-term work on plantations, with the Australasian Petroleum Company (APC) and with the government. During this time he also learned some 'pure' Motu and more recently has been trying to learn some Tok Pisin (the official name for what used to be called New Guinea Pidgin (English), Neo-Melanesian, and Tok Boi).

There are some points about Nanai's background that help explain the origin of the unexpected English-based items in his Police Motu:

- (a) he learned to speak English reasonably well and does not seem to have gone through a 'broken English' stage characterised by the use of long, bilong, verbs marked by -im, and adjectives marked by -pela, at least this is not evident in his present-day speech as is indicated by the sample reproduced here in Appendix 5;
- (b) he did not know any Tok Pisin at the time he made the recording and has not learned much since;
- (c) he learned Police Motu before the war.

For these reasons we can be as sure as it is possible to be without more extensive records and earlier comparative materials that at least the following, if not all, of the above words were part of Police Motu before the war: *mistim*, *nambawan*, *seken*, *stanibai*, *traim*, *halusia*, and that the remainder (except for *et thausen* whose value no one at that time could possibly have appreciated) were probably so, or were coming in. This is so particularly because the set of forms in the 'sure' subset above (i.e. *mistim*, *nambawan*, etc.) are non-standard English ones in some way (e.g. *halusia*) and/or are common to Pidgin English of the south-west Pacific (*mistim*, *nambawan*, *seken*, *stanibai*, *traim*). Indeed the most likely source of these forms was the 'broken English' spoken throughout most of Papua before the war and introduced to eastern and western Papua before Police Motu was ever heard of.

3.2.4.2 Other evidence

As a check on these conclusions, however, and in an attempt to more clearly define the nature of Police Motu before the war, a survey was made of other speakers (both Papuan and expatriate) of Police Motu who had learned their Police Motu before the war.²⁵ The following techniques were used:

- (a) Nanai's recording was played to speakers of Police Motu from other parts of Papua but especially from the Western Province since, according to the 1961 survey of Police Motu (Brett et al 1962:11), the form of Police Motu spoken there is intelligible to all other Police Motu speakers. It was

from this area also that recruits were mostly drawn in the early days of the first official police force. These were then asked to repeat those sections containing the above list of words and to explain these in their own words;

- (b) a list of words containing those items which I as a speaker of Police Motu since 1957 thought were "incorrect" or suspicious as being forced borrowings from Motu (forced in the sense of being there because the recorder thought they should be) was compiled from the three wartime published "descriptions" discussed above. This list is given in Appendix 6 and was used as a base to survey the content of the Police Motu of pre-war learners of the language;
- (c) speakers were questioned about vocabulary used in occupations that were common before the war (e.g. plantation, marine and mining industries) to try to fill out the vocabulary of possible English borrowings then current;
- (d) speakers were observed when speaking more casually and any interesting items noted.

The results of this investigation allow the following conclusions to be drawn:

- (1) Nanai's speech was typical of the time and readily understood by other speakers (and presumably this is why he was originally chosen to make this propaganda recording);
- (2) the English-based words in Nanai's speech are never given their "proper [English] pronunciation" as the author of the 1941 Handbook had urged earlier but are Papuanised or given a pronunciation in accord with the phonotactics of speakers of Papuan languages. Nor are they always recognised as English borrowings. Indeed many have become so integrated that they are not recognised as English borrowings but are regarded as part of Police Motu;
- (3) speakers recognise that synonyms exist for many items (e.g. mamoe and sipi for *sheep*, hanua and tauni for *town* - see Appendix 7) but choice of which item to use when, generally depends on other factors (e.g. the company one is in at the time of speaking);
- (4) many speakers can vary their speech along a scale towards Motu and what is today called the Central dialect of Police Motu because of their contact with Motu speakers in Port Moresby, or on government stations where Motu speakers were engaged as clerks in the days when they were better educated than other Papuans, or on London Missionary Society mission stations;
- (5) English and Motu were not the only sources of borrowing into Police Motu. Some items were borrowed from individual Papuan languages, but especially Suau, the Austronesian language spoken along the southern coast just west of Milne Bay (see Appendix 8). Thus the items kamkam *fowl*, bush *fowl*, 0 kapore *oh sorry*, susu *milk*, and dimdim *white man* are to be traced to this source.²⁶ Others have come from Koriki (*nakimi brother-in-law*), from one or more Binanderean languages of the Northern Province (*kiki to yam*),²⁷ and from languages in the Torres Straits (*mamoos*, *mamus chief, policeman*);
- (6) there were and still are many more English-based words in Police Motu than is generally recognised, and this number has been constantly increasing as new cultural items have been introduced into the country (as was recognised by the author of the 1941 Handbook as noted above). Some examples are given in Appendix 7;
- (7) the linguistic picture was even more complicated than the 1961 survey (Brett et al 1962) indicates. Thus there was variation between areas (see point 5

above), across speakers (point 4 above), and according to company (see point 3 above).

- (8) the 1941 Handbook materials are closest to Nanai's in phonology and vocabulary but it is clear, as suspected, that some of the vocabulary presented in that account are elicitation or learning errors. On the other hand many of the Motu words in the vocabularies of the other two accounts were not part of Police Motu, at least as spoken by those outside of the Port Moresby area - see Appendices 7 and 8;
- (9) the following grammatical features discussed earlier were not part of Police Motu spoken outside of the Port Moresby area:
 - (i) -ai as direction or location marker;
 - (ii) baine ... neganai *if*;
 - (iii) edeheto *how*;
 - (iv) object suffixes on verbs.

4. CONCLUSION: THE NATURE OF POLICE MOTU BEFORE THE WAR AND RELATED ISSUES

Now that all relevant materials and evidence have been presented and discussed it is possible to return to some of the questions raised by the wartime written and oral material discussed above and in the process to say something more precise about the nature of Police Motu before the war and subsequently.

The following seem to be the most pertinent observations to make:

- (1) pre-war Police Motu was very similar to post-war and modern Police Motu. In particular it probably consisted of two dialects as today but with complex variation within and between these, particularly in respect of phonology and vocabulary, but also to some extent (though not very much) in grammar;
- (2) pre-war Police Motu was never purely Motu-derived in vocabulary but consisted of a basic vocabulary drawn from Motu (generally for those items and concepts that were part of Papuan culture at the time of contact, but not completely) a few loans from other languages of Papua, and an ever-increasing number of English-based loans (that kept pace with introduced cultural items and ideas). In the early days "broken English" was the principal source of these loans but gradually this gave way to standard English as the source (and now to Tok Pisin as well - for example, pamuk *prostitute* is replacing siari kekeni) as knowledge of English improves. In the 1941 Handbook this phenomenon was recognised although the implications of it were not followed through. The myth that Police Motu consists solely of Motu-based vocabulary (except for a few well-known exceptions) continues to be perpetuated in written descriptions of the language today. However, once the above fact is appreciated Lock's statements noted above in section 3.1.2. about Police Motu being difficult for Motu speakers to understand and about policemen adding more Motu words to the bastard language become intelligible or take on a new significance. These would be natural consequences of a language full of unfamiliar English and other language-based loans. By the same token this also probably partly explains why Police Motu is so similar to present-day Tok Pisin in many respects, a question raised some time ago but never adequately answered (Dutton 1976);
- (3) there was never any standard form of Police Motu but the tendency has always been to regard that form of it that is 'closest' to Motu as the standard

form. This tendency is clear in two out of the three sets of written material described herein and is continued in all but the latest written descriptions of the language after the war. Indeed in the earliest and best known of these accounts (Chatterton 1946, 1950, n.d.) an attempt was made to tidy up the language and to create an improved version which "while maintaining a reasonable measure of simplicity" avoids "unnecessary barbarities" (Chatterton 1950:5). This tendency to try and upgrade the language by ('pure') Motu-ising it was bound up with European notions of correctness as well as with their negative attitudes towards pidgin languages in general at the time. In Police Motu's case the language was constantly compared to 'pure', or (as Sir Hubert Murray, the first and longest serving Lieutenant Governor of Papua, insisted on calling it) "classical" Motu, and its "barbarities" emphasised;

- (4) the war had a very significant impact on Papuans and on their lingua franca Police Motu. Because of the large numbers of Papuans involved as labourers in the war effort and because of the language policies and practices adopted by the allied forces Police Motu was given a significant boost in status: it had been formally recognised as an important language whose use had been actively encouraged by powerful bosses, it had been described and recorded, and it had been used in broadcasting for the first time. But the language itself does not appear to have been affected much by this change except that many new English-derived war-related vocabulary items were added to its lexical store.

NOTES

1. This paper is the result of an ongoing investigation into the history of Police Motu, the full results of which will be published in book form, tentatively titled *Police Motu: iena sivarai*. I am grateful to the Australian National University for the funds and opportunity made available to carry out the research upon which this paper is based. I am also grateful to the many colleagues who have listened to or read other versions or parts of the details presented in this paper over the past year or so, although none is to be held responsible for the views presented herein.
2. Before the war the country that is today called Papua New Guinea consisted of two separate territories, the Territory of Papua and the Trust Territory of New Guinea. During the war a military administration replaced the civilian one. During this period both territories were officially called "Australian New Guinea". After the war the two territories were combined as Papua-New Guinea until 1949 when they became the Territory of Papua and New Guinea (Nelson 1982:fn.1).
3. For fuller details on aspects of the war in the Pacific and Papua New Guinea in particular see Barrett 1969, Long 1973, McCarthy 1959 and Nelson 1982.
4. The other was 'broken' English which is referred to in more detail later.
5. Although the language is officially known as Hiri Motu, and has been since 1970 when the name was first mooted, it is not generally used by speakers outside of the Port Moresby area. Most of those I have questioned about this name confuse it with 'pure' Motu because of the widely known association

of the Motu with the *hiri* (the Motu name for their (former) annual trading voyages to the Gulf of Papua).

6. F.E. Williams, the Government Anthropologist at the time, described it as such in an unpublished manuscript giving general hints about how army personnel unfamiliar with Papuans and New Guineans should treat them (Australian War Memorial, File 506/1/1).
7. It is difficult to be precise about this because no records were kept. But there are sufficient references to the use of "English" in the force and attempts to improve it that it is fairly clear what the situation was.
8. This description of the distribution of Police Motu before the war is based on statements in the wartime written materials discussed in this paper and on the distribution statements made in the published account of a linguistic survey of Police Motu made in 1961 (Brett et al 1962) referred to later in this paper.
9. Unbelievable as this may sound I have so far been unable to locate a single note, letter, or other material in the language before the war. Part of the reason for this is probably that the language was mainly a spoken one and few letters or records of it were kept; another part of the reason is probably that no one ever thought about preserving material in it - it was after all regarded as hardly a language in European eyes.
10. The background to this is that in early 1941 Allied Intelligence was informed that Japan seemed likely to enter the war soon to assist Germany. If so it would move southwards towards the British stronghold of Singapore as a first and immediate goal (Long 1973:107-108). 1941 therefore saw a quickening of preparations for war in Australia. As part of those preparations additional troops were sent to Port Moresby and intelligence gathering in the Australian region stepped up (McCarthy 1959:9-12;41).
11. These are located in the Australian War Memorial in the following places: (a) File 5.6; (b) File 506/9/4; and (c) File 506/9/1. I should like to thank the Australian War Memorial for permission to carry out research in their archives and for the patient assistance provided in helping to locate some of this material during renovations to the Memorial.
12. I have been unable to identify the authors of these two manuscripts.
13. As will be pointed out below Pastor Lock was subsequently killed in 1951 when Mount Lamington erupted, so that he cannot be questioned about his manuscript. Information about his life and work and opinions about his manuscript have, however, been obtained from his brother Pastor L.N. Lock, who also lived and worked in Papua and knows, and still does translation work in, *Hiri* (or *Police*) *Motu*. I should like to thank Pastor Lock for his generous assistance in this regard and for kindly letting me hear a recording of his brother making a radio broadcast after the war. This information is used in assessing Maynard's material below in section 3.1.2.
14. I use my own knowledge of the language as a guide in making decisions about what *Police Motu* was like after the war as I first went to Papua in 1957 and learned *Police Motu* soon afterwards. There are also good comparative texts available in my own and colleagues' fieldwork tapes from 1965 onwards, some of which appear in Dutton and Voorhoeve 1974 for example.
15. Note that this and the following statements are only generally true. There are counter examples to each of the following points but only sufficient to prove the rule. The author also makes mistakes sometimes, e.g. when he/she writes tau for *man* as in *Motu*, when by the author's own rules (and assuming

- that the word has not changed form in the meantime) this should have been written tow. The author also uses double consonants at times (e.g. gar-ree for gari *afraid*) but these are ignored for present purposes as they are not consistent and do not seem to indicate particular phonological features.
16. In these and other examples the Handbook entries are respelled using standard Police Motu orthography, that is, the sounds represented by the English orthography are given their Police Motu orthographic values. Note that this is not the same as giving them the standard Police Motu form.
 17. This claim is based on my own extensive experience of the pronunciation of Police Motu by speakers from different areas of Papua after the war. It could be formally documented, however, if necessary, but some support for it can be obtained by listening to appropriate texts in Dutton and Voorhoeve 1974:Units 11 and 12.
 18. For example, Sir Hubert Murray who was Lieutenant Governor of Papua for 33 years regarded the language as "a bastard jargon, almost as bad as pidgin English" and hoped that it would "also, in time, disappear" (Murray 1930: iii).
 19. The following details come from Australian War Memorial File 431/8/3: "Propaganda Tour by New Guinea Natives". This particular quotation is taken from p.2 of this document.
 20. This was Sergeant Pegano of Manus Island, as indicated by the text in Pidgin English on the end of the recording containing the Police Motu text.
 21. This party was under the official care of Warrant Officer "Bill" Gordon of ANGAU. It was also accompanied by at least one other European, Mr G.A.V. Stanley, a geologist in pre-war Papua fondly referred to by Papuans as *uda boroma bush pig*.
 22. A copy of the transcription is also being made available to the archives of the Australian War Memorial and, it is hoped, will be published later.
 23. This is easily confirmed by listening to Nanai's parallel propaganda text in Kerewo already referred to.
 24. Other suffixes do not appear in the text because of the nature of the text but when interviewed recently about the use of other suffixes (e.g. -gu *me*, -mu *you*, etc.) Nanai explained in essence that he does not use them unless he is in Port Moresby or is speaking to a Motu speaker.
 25. The opportunity was also taken to ask speakers about the distribution, use and understanding of the official name Hiri Motu whenever possible as well as about the present situation of the language. The results of this brief investigation will be reported elsewhere.
 26. Susu is presumed to have come directly from Suau although it may well have come from 'broken English' as it is a common form for *milk* in Pidgin Englishes of the south-west Pacific.
 27. The generally hitherto accepted view of the origin of kiki is that it comes from the Korafe language near Tufi. However, it is more likely that it comes from one of the other Binanderean languages, most likely Binandere itself. But as investigations into this are still continuing it is not possible to give a more definite answer to the question here.

Appendix 1: Excerpt from the 1941 Handbook of Motuan (Police Motu)

NOTES ON MOTUAN

The following vocabulary, by no means comprehensive, is made up of words and phrases in common use and the spelling is as near as possible to the pronunciation of the word.

This language is used universally in the Central Division of Papua and east along the coast until the influence of Samarai is reached. Although understood by many in this area, it is not generally used, Soo-ow and English replacing it. Its use is quite general along the coast west of Port Moresby, until it gives way to the influence of Bamu and Kiwai tongues. Even here it is not out of place. The language is quickly acquired by natives who have been recruited for service in the districts under the influence of Port Moresby, and they retain a good knowledge long after they return to their homes. This has brought about the universal use of the language.

When speaking Motuan, any tendency to drawl must be avoided, as it is normally spoken quickly, and slowness will cause it to be misunderstood. Intonation plays a big part, and should be practised from the start.

It will be seen by a perusal of the Vocabulary, that quite a number of words when used to refer to the first person have *gu* added to the stem, e.g.:

ENGLISH	MOTUAN EQUIVALENT	1st PERSON
<i>Father</i>	tama-na	tama-gu
<i>Mother</i>	sina-na	sina-gu
<i>Give</i>	henni-a	henni-gu

The *gu* is derived from:

ENGLISH	MOTUAN
<i>Me or I</i>	lau
<i>My or mine</i>	lau-egu

Except where preceded by the pronoun, the noun is always used first, e.g.:

ENGLISH	MOTUAN
<i>Bring the axe</i>	ee-la mai-laya
<i>Bring your axe</i>	oi-emu ee-la mai-laya

Quite a large proportion of the natives who are in constant contact with the white man speak good English and in dealing with these the stranger has no difficulty. A native will more readily understand good English than Pidgin, the use of which is always discouraged.

If you cannot find a word in the Vocabulary such as horse, launch, bicycle, motor-car, use the English word with its proper pronunciation. All such importations of civilisation have no word in the Motuan tongue and English is adopted.

Under Native Regulation, No.1 of 1907, the attendance of native children at school is compulsory, providing English is taught. It appears from the reports of the magistrates and others that English is now taught at practically all mission stations where there are white teachers.

By an order-in-council, white officials are instructed to use English in speaking to Government native employees, police, and prisoners, and to extend, where possible, a knowledge of English.

CARDINAL NUMBERS

<i>one</i>	tah
<i>two</i>	rua
<i>three</i>	toi
<i>four</i>	har-nee
<i>five</i>	ee-ma
<i>six</i>	taura-toi
<i>seven</i>	heetu
<i>eight</i>	taura-harnee
<i>nine</i>	taura-harnee-ta
<i>ten</i>	ku-ota

ENGLISH	MOTUAN (as pronounced)	ENGLISH	MOTUAN (as pronounced)
A		B	
<i>about (to go about)</i>	giroa-low (ow as in plow)	<i>baby</i>	nar-toona
<i>above</i>	a-taya	<i>bad</i>	deeka
<i>a bit of</i>	taina	<i>bag (net)</i>	kee-arpa
<i>accept</i>	ubia (ub as in hub)	<i>banana</i>	bee-koo
<i>a great many</i>	mor-mor herea	<i>barber</i>	wee-wheena
<i>afraid</i>	gar-ree	<i>bandicoot</i>	ootooa towna
<i>afraid of</i>	gar-ree gar-reena	<i>bark (of a tree)</i>	mar-da
<i>again</i>	nega-tar (sound e as air)	<i>basket</i>	ow-korpeena
<i>age, old</i>	buruka	<i>batchelor</i>	kee-arpa
<i>aid me</i>	durugu	<i>bathe (place to)</i>	oo-ow towna
<i>aid (others)</i>	durua	<i>bath</i>	dee-goo
<i>afternoon</i>	ador-ray	<i>bathroom</i>	dee-goo ga-boona
<i>a little</i>	tai-na, or sisina	<i>battle</i>	dee-goo gowna
<i>all</i>	ee-bone-ay	<i>beach</i>	dee-goo roomah
<i>alongside</i>	buddi-budina (i as ee)	<i>beat (to strike)</i>	too-warry
<i>almost (nearly)</i>	kyla-kyla	<i>beats</i>	kornee
<i>altogether (at one time)</i>	nega tar-mona	<i>because</i>	bot-taya
<i>also</i>	dar-noo	<i>bed</i>	ar-gewa (e sound as air)
<i>always</i>	hunna-hunay	<i>bedroom</i>	ba-dinna
<i>angry</i>	daggi-daggi	<i>behind</i>	mar-hoota gowna
<i>ankle</i>	hai-na	<i>bell</i>	mar-hoota gubuna
<i>arise</i>	torisi (i as ee)	<i>bell (to ring)</i>	muri-nay
<i>arm</i>	ee-man-a	<i>belly</i>	gar-ba
<i>armlet (plaited)</i>	gar-na	<i>below (in the valley)</i>	gar-ba tor-wa
<i>ask</i>	h'nana-daya	<i>beneath (under)</i>	bo-kana
<i>assault</i>	bot-taya	<i>beg (or cadge)</i>	hen-u-henu-ay
<i>assist</i>	duru-a	<i>big (large)</i>	hen-u-nay
<i>at once (quickly)</i>	ka-hurra-ga	<i>bind (tie up)</i>	noi-e-noi-e
<i>awake (to)</i>	torisi (i as ee)	<i>bird</i>	barda herea
<i>axe</i>	eela	<i>biscuit</i>	kwut-tooa
		<i>bite</i>	marnoo
			kis-kis
			kolia

ENGLISH	MOTUAN	ENGLISH	MOTUAN
<i>bitter (unripe fruit)</i>	ku-see-ree	<i>cook (to)</i>	nar-dua
<i>black</i>	koremi-koremi	<i>cost</i>	davana
<i>bleeding</i>	lala-la dio	<i>cost (what is cost of)</i>	davana heda
<i>blind (cannot see)</i>	mata-keboori	<i>count</i>	doo-waya
<i>blood</i>	lar-la	<i>crocodile</i>	oo-wulla
<i>boiling water</i>	wasee-oh	<i>crooked (not straight)</i>	ga-gewa
<i>boots</i>	tamukka	<i>cry</i>	tai
<i>bone</i>	too-re-enna	<i>cut</i>	oo-tooa
<i>bottle</i>	boh-tohl	<i>cut (firewood)</i>	lah-hee siria
<i>bow (and arrow)</i>	pow		
<i>bowels</i>	bo-kana	D	
<i>boy</i>	merro	<i>damp</i>	purri-purri
<i>boy, small</i>	merro mara-kee	<i>dance</i>	ma-var-ru
<i>box</i>	mow-wa	<i>dark</i>	di-boora
<i>bread</i>	pa-lowa	<i>dawn</i>	dubba-raeri
<i>break</i>	makoi-ia	<i>day</i>	dinana
<i>bridge</i>	oo-wupoo	<i>day (this)</i>	hari-dinana
<i>bring</i>	my-laya	<i>dear (expensive)</i>	davana-barda
<i>bring more</i>	haida-mai-laya	<i>deep</i>	dobu-dobu
<i>broom</i>	dar-hua gowna	<i>die (or dead)</i>	massi
<i>brother (elder)</i>	kukunna	<i>different</i>	eedau
<i>brother (younger)</i>	tadina	<i>dig</i>	gay-ia
<i>build or make</i>	gurria	<i>dinner-time (midday)</i>	dina-toopoo
<i>burn</i>	garboo-a	<i>dirt (on anything)</i>	mee-ru
<i>bury</i>	hardoa-tarnoo-ai	<i>dirty</i>	mee-ru mor-moa
<i>buy</i>	hoi-ia	<i>distant</i>	dow-dow
C		<i>distant (very far)</i>	dow-dow herea
<i>call</i>	boi-boi	<i>ditch, drain or stream</i>	sina-wy
<i>cane (Loia)</i>	ow-ru	<i>do</i>	kurria
<i>canoe</i>	wonnugi	<i>do you understand</i>	oi-deepa
<i>canvas</i>	pa-lai	<i>dog</i>	siss-ia
<i>cap</i>	kora-gowna	<i>door</i>	eedoo-wurra
<i>careful (to be)</i>	nah-ria namo namo	<i>doubtful</i>	sed-ila
<i>carry</i>	hoo-waya	<i>dream</i>	nee-hee
<i>case (box)</i>	mow-wa	<i>drink</i>	ee-noo-a
<i>chair</i>	helai-gowna	<i>drum</i>	garba
<i>chase</i>	roo-rooa	<i>drunk</i>	kava (mura-mura kava)
<i>chicken</i>	koka-rogo	<i>dry</i>	kow-kow
<i>chicken (Samarai and District)</i>	kum-kum	<i>dugong (or sea cow)</i>	loo-y
<i>child</i>	nar-toona	E	
<i>clay</i>	tar-noo	<i>ears</i>	ty-una
<i>climb</i>	dy-kow	<i>earth</i>	tar-noo
<i>clock</i>	dina-gowna	<i>eat</i>	anna-ia
<i>closet (W.C.)</i>	kukuri-gaboona	<i>eggs</i>	kut-toy
<i>cloth, clothes</i>	darboo-a	<i>eight</i>	towra-harni
<i>coconut</i>	ni-u	<i>embrace</i>	ross-ia
<i>cold</i>	keru (e sound as air)	<i>empty</i>	anina-las
<i>comb</i>	edu-wurri	<i>enough</i>	wadarni
<i>come</i>	mai		

Appendix 2: Excerpt from Lock's "Police Motu"

"Police" Motu is the principal language spoken almost throughout the Territory of Papua. "Pure" Motu is the language spoken by the coastal natives in the MORESBY District, or as it was originally known The Central District. This area extends on the East as far as GABAGABA and on the West as far as MANUMANU.

Strange as it may seem, "Police" Motu is difficult to understand by many "Pure" Motu speakers, unless they have had much practice. It sounds as bad to them as "Pigin" English does to us, and it is really a bastard language.

Nevertheless, to the newcomer, it is a much easier language to learn than "Pure" Motu. The grammar is much more simple, and there are very few set rules. It is called "Police" Motu because it was originally made up and spread throughout the Territory by Native Police, who, after being trained in Port Moresby, were sent to outstations in areas of a different language group. Each picked up more Motu words and added them to the Bastard language.

The main difference lies in the noun and the pronoun, the latter usually taking the third person. In verbs, there is no change in tense, as in "Pure" Motu. There are also other minor differences which are dealt with under separate headings.

MOTU GRAMMAR

In Police Motu there are no set rules of grammar. The following headings set out practically all that is necessary to speak the language.

The alphabet consists of 19 letters: a, e, i, o, u, b, d, g and g, h, k, l, m, n, k, p, s, t, v, and two compound letters, kw and gw.

The vowels have the Continental sound and are usually long as in: father, met, it, more, put.

The Consonants are pronounced as in English but at times the g has the ng sound.

There are many Diphthongs, e.g.: ae, ai, ao, au, ei, eu, oi, and ou. No two consonants ever stand together.

There is no rule for accent.

1. THE NOUN. Nouns are usually primitive as: au *tree*, nadi *stone*.
2. GENDER. This is not shown other than by the meaning of the word.
3. NUMBER. This also does not generally affect the noun. Hisiu is *star* or *stars*. A very few nouns have the first syllable duplicated for the plural form: mero *boy*, memero *boys*, tau *man*, tatau *men*.
4. The PRONOUN. Personal pronouns only have the person.

Singular			Plural		
1st	lau	<i>I</i>	1st	ita	<i>we(incl.) ai (excl.)</i>
2nd	oi	<i>thou, you</i>	2nd	umul	<i>you</i>
3rd	ia	<i>he, she, it</i>	3rd	idia	<i>they</i>

POLICE MOTU SYNTAX

THE SIMPLE SENTENCE. In a Police Motu sentence the subject or nominative usually stands first, followed by the Object, or accusative, if any. The Verb generally goes at the end of the sentence. For example:

Unai mero maragi lauegu tadina ia botaiia.
That little boy hit my young brother.

Lauegu imana lau utua.
I cut my hand.

Unai ira mailaia.
Bring that axe.

Unai be daika ena natuna?
Whose child is that?

COMPLEX SENTENCES. In a complex sentence the Object still comes last. e.g.:

Ruma haginia tauna ia heau vadaeni.
The man who built the house has run away.

Ai lao unai motumotu dekenai, vadaeni kerukeru ai giroa lou.
We are going to that island, and will return tomorrow.

It is hard to set any rule in the clauses of the sentences, as it would be of no practical use, as nearly all the different Tribes have their own formation of the Police Motu sentences. As I have already stated it is not a real language and the best I can do is to offer a few examples of the general formation of the sentences.

1. Will be the Motu sentence.
 2. Will be the English translation.
 3. Will be the English meaning.
1. Dohore hari dina, lau lao digudigu.
 2. *Yet now day I go swim.*
 3. *I am going for a swim today.*
1. Oi dekenai aniani haida ia noho?
 2. *You with some food it stop?*
 3. *Have you got anything to eat?*
1. Hari dina be edeseni oi lao?
 2. *Now day is where you go?*
 3. *Where are you going today?*
1. Lauegu tamana be edeseni?
 2. *My father is where?*
 3. *Where is my father?*
1. Lauegu kakana be kahirakahira ia mase.
 2. *My brother is nearly he dead.*
 3. *My brother is dying.*
1. Motuka dekenai ia moru tauna, harihari sisina ia namo.
 2. *Motor-car at he fell man, now bit he good.*
 3. *The man who fell off of the truck is now getting better.*

<i>word</i>	hereva	<i>yam</i>	maho	<i>yonder</i>	ununseni
<i>work</i>	gaukara	<i>yawn</i>	mava	<i>you (sg.)</i>	oi
<i>world</i>	tanobada	<i>year</i>	lagani	<i>you (pl.)</i>	umui
<i>worship</i>	guriguri	<i>yell</i>	boiboi	<i>young</i>	natuna, karukaru
<i>write</i>	torea	<i>yes</i>	oibe, io	<i>youth</i>	tauhu
<i>wrong</i>	kerere	<i>yesterday</i>	varani		

PHRASES

Lau mai.	<i>I come. I am coming. I have come.</i>
Oi mai.	<i>You come. You are coming. Are you coming?</i>
Ia mai.	<i>He comes. He is coming. Is he coming?</i>
Lau gorere.	<i>I am sick.</i>
Oi gorere?	<i>Are you sick?</i>
Ia gorere.	<i>he is sick.</i>

COMPLEX SENTENCES

- The aeroplane fell down in the hill over there and I want you to go and find it.*
- Unai aeroplane be ororo kahanai ia moru diho, lau ura oi lao bona oi tahua.
- When I tell you to do something I don't want you to stand there and look at me, go and do it.*
- Baine lau hamaoromu neganai gau ta oi kataia, lau ura lasi unuseni oi gini bona oi itagu, oi lao bona oi karaia.
- Today we are going for a long walk, and I do not want any of you to sit down on the way, because this is a test of your strength.*
- Hari dina ita lao loaloa daudau, vadaeni lau ura lasi ta daladekena, umui helai diho badina inai be umui goada ai tohoa.

NUMERALS

<i>one</i>	ta, taomona	<i>twenty</i>	ruahui
<i>two</i>	rua	<i>thirty</i>	toihui
<i>three</i>	toi	<i>forty</i>	haniahui
<i>four</i>	hani	<i>fifty</i>	imahui
<i>five</i>	ima	<i>sixty</i>	tauratoi-ahui
<i>six</i>	tauratoi	<i>seventy</i>	hitu-ahui
<i>seven</i>	hitu	<i>eighty</i>	taurahani-ahui
<i>eight</i>	taurahani	<i>ninety</i>	taurahani-ta-ahui
<i>nine</i>	taurahani-ta	<i>hundred</i>	sinahu
<i>ten</i>	gwauta	<i>thousand</i>	daha

Note: after every ten first numeral is repeated such as *thirteen gwauta-toi*, *twenty-three ruahui-toi*, *forty-three hariahui*, and so on.

COLOURS

<i>black</i>	koremakorema
<i>blue</i>	gadoka-gadoka
<i>brown</i>	gagare
<i>red</i>	kakakaka
<i>green</i>	kasiri-kasiri

aba; s.	<i>bush turkey</i>	abia lou	<i>take again</i>
abata	<i>flooded river</i>	abia mai	<i>to bring</i>
abia; v.t.	<i>to have; to hold</i>	abo	<i>scrotum</i>
abiahanaia	<i>take across</i>	advana	<i>husband, wife</i>
abia hebou	<i>take together</i>	adavaia	<i>to marry</i>
abiaisi	<i>to lift up</i>	adorahi	<i>afternoon</i>
abia kava	<i>take without reason</i>	ageva	<i>beads</i>
abia lasi	<i>do not take</i>	ahetoni	<i>parting with friends</i>

ahu	<i>lime</i>	au ramuna	<i>roots</i>
ai	<i>we (excl. of person addressed)</i>	au rigina	<i>branches</i>
aiha	<i>centipede</i>	auka	<i>firm, hard, difficult</i>
aivara	<i>canoe pole</i>	auka masemase	<i>very hard</i>
ane	<i>song, hymn</i>	auri	<i>iron, metal</i>
aneru	<i>angel</i>	ava	<i>weeds</i>
ani	<i>used at end of sentence when answer expected</i>	ava raua	<i>to weed</i>
		B	
aniani	<i>food, to eat</i>	bada; a.	<i>large, gread</i>
anina; s.	<i>flesh of body</i>	s.	<i>size</i>
anina	<i>substance; meaning of a thing</i>	badina	<i>equivalent (because)</i>
aonega; s.	<i>wisdom, intelligence</i>	badina; pp.	<i>by the side of, nearby</i>
a.	<i>wise</i>	badu	<i>angry, anger, to be angry</i>
araia	<i>to set fire to</i>	bamona	<i>the same, companion</i>
aria	<i>feast</i>	bara; s.	<i>oar</i>
ariara natuna	<i>illegitimate child</i>	v.i.	<i>to row</i>
ariara hahine	<i>prostitute</i>	baragi; s.	<i>lungs</i>
asimana	<i>to sneeze</i>	baubau	<i>bamboo</i>
atai ai	<i>above</i>	bava	<i>a crab</i>
atoa	<i>to place, set</i>	be	<i>used in a question with the meaning of why?</i>
au	<i>tree</i>	bema	<i>with</i>
au badina	<i>trunk of a tree</i>	beni	<i>spleen</i>
au burena	<i>flower of a tree</i>	bero	<i>wound</i>
au dorina	<i>top of a tree</i>	bese	<i>nation</i>
au kopina	<i>bark of a tree</i>	biaguna	<i>master, owner</i>
au maragina	<i>small tree</i>		

Appendix 3: Excerpt from the "Police Motuan Vocabulary"

VOCABULARY

ENGLISH	POLICE MOTUAN	ENGLISH	POLICE MOTUAN
(A)			
<i>abdomen, belly</i>	bogana	<i>box</i>	maua
<i>above</i>	atai	<i>boy</i>	mero
<i>ache</i>	hisisi	<i>break (in half)</i>	kwaidu
<i>across (to go)</i>	hanai	<i>break (smash)</i>	makohi
<i>adze</i>	omo	<i>bring</i>	mailaia
<i>afraid</i>	gari	<i>brother: younger</i>	tadina
<i>another kind</i>	idau	<i>elder</i>	kakana
<i>afternoon</i>	adorahi	<i>burn</i>	gabua
<i>aid or assist</i>	duruu	<i>bury</i>	guria
<i>alive</i>	mauri	<i>buy</i>	hoihoi (hoia)
<i>all</i>	ibounai	<i>blood</i>	rara
<i>alone</i>	sibona	(C)	
<i>I alone</i>	lau sibogu	<i>call</i>	boiboi
<i>he alone</i>	ia sibona	<i>canoe</i>	vanagi
<i>also</i>	danu	<i>carry</i>	huaia
<i>and</i>	bona	<i>carry away</i>	laohuaia
<i>altogether</i>	ibonai	<i>catch</i>	gobia
<i>anger</i>	badu	<i>chest (of man)</i>	keme
<i>another</i>	ta	<i>child</i>	natuna
<i>any</i>	haida	<i>chop</i>	siria
<i>arise</i>	toreisi	<i>close</i>	kailakaila
<i>arm</i>	imana	<i>cloth</i>	dabua
<i>arrive</i>	ginidai	<i>cocconut</i>	niu
<i>ask</i>	henanadai	<i>cold</i>	keruma
<i>assist</i>	duruu	<i>come</i>	mai
<i>auger</i>	ibudu	<i>cook</i>	nadua
<i>axe</i>	ira	<i>copper</i>	kepa
		<i>crab: large</i>	bava
		<i>small</i>	dubara
(B)		<i>crayfish</i>	ura
<i>beg</i>	noinoi	<i>crocodile</i>	huala
<i>bad</i>	dika	<i>cry</i>	tai
<i>bag (rice)</i>	pusi	<i>cut</i>	utua
<i>bag (native)</i>	kiapa	<i>cuscus (o'possum)</i>	vaura
<i>bamboo</i>	baubau	(D)	
<i>banana</i>	biku	<i>damage (to)</i>	hadikaia
<i>basket</i>	bosea (bosia)	<i>dance</i>	mavaru
<i>bathe</i>	digu	<i>day</i>	dina
<i>beach</i>	kone	<i>dead</i>	mase
<i>beads</i>	ageva	<i>deaf</i>	taia kudima
<i>because</i>	badina (used also as <i>why</i>)	<i>deep</i>	dobu
<i>behind</i>	murinai	<i>dig</i>	geia
<i>bell</i>	baga	<i>dirt</i>	miro
<i>beyond</i>	unukahana (<i>the other side</i>)	<i>dog</i>	sisia
<i>big</i>	bada	<i>door</i>	lduara
<i>black</i>	koremakorema (<i>some- times korema</i>)	<i>draw water</i>	ranu utua
<i>blind</i>	matakepulu	<i>dream</i>	nih
		<i>drink</i>	inua
		<i>drum</i>	gaba

*Where an obvious typographical error has occurred in this listing the correct form is given in []s following an 'equals' sign.

SENTENCES

OF BUILDING

We will build a house here for the boys to sleep in

Take axes and knives and go into the bush and cut the necessary timber

Cut plenty of long poles

Some boys will cut grass for thatching

Don't cut the short grass

Only long grass is of any use

Tie these poles securely

Dig deep holes for the uprights

Ram the earth well down beside the post

If the post moves when pushed it is no good

Stand the post straight up

Tie on plenty of scaffolding

The rain may come through the thatching in several places, close them up securely

Make a door at this end

Too many boys are splitting cane. Some are to come and assist the builders

OF THE SEA

The canoe is on the beach

The paddles are on the canoe

The mast and sail are also there

The sea is very rough

We will wait for high tide

The canoe has sunk

Some boys go and fix it up (done by rocking back and forth)

You go and catch fish (with hook and line)

You go and catch fish (with spear)

MEDICAL

Are you ill?

I have a bad cough

I have a head-ache

Memerodia mahuta rumana ita karaia
inesini

Ira bona kaia abia lao udai ruma karaia
auna utua

Au latadia momo utua

Memero haida idia kurukuru utua

Kurukuru kwadogina utua lasi

Kurukuru latana sibona ia namo

Ine au kwatua namu namu

Augini gurina geia dobu dobu (ha'dobua)

Tano gwadaia auka, au badibadina

Bema au doria ia marere-marere vadani
ia dika

Au haginia maoro maoro

Aena moia auna momo kwatua

Gabu haida medu ia diho garina, kohua
namo namo

Ia kahana eduara karaia

Memero momo idia waro haparara-ia a
haida mai ruma karaia taudia durua

Vanagi ia noho kone-ai

Bara idia nono [=noho] vanagi dekena

Autubu bena lara danu idia noho vanagi ai

Davara ia hurehure mo mo

Davara bada ita naria

Vanagi ia mutu

Memero haida idia lao, pataia

Oi lao kwarume veria

Oi lao kwarume gwadaia

Oi gorere bada a?

Lau huahua dikadika

Laegu kworana ia isis i

*What is the matter with you?
 How did you cut your leg?
 The swelling came of its own accord.
 Have those sores on your face been
 treated?
 I have the symptoms of dysentery
 (there is blood in my excreta)
 I have pains in the stomach
 The cough causes a pain in my chest*

GENERAL

*You come here
 Come up here
 Go down there
 Listen to what I have to say
 Do you understand?
 When will he arrive?
 Come quickly when I call
 Did he send a letter (note)?
 I will send a letter (note) later
 Why have the boys been quarrelling?
 One boy has been stealing
 One boy has stolen something
 The boys have gone hunting
 Take this ... away
 Place it on the store verandah
 Bring me that box*

THE TRACK - CARRYING

*We leave at daylight tomorrow morning
 How many nights must we spend on this
 track?
 How many large hills are there to
 cross?
 How many rivers are there on this
 track?
 Are there any big rivers?
 Are any likely to be in flood?
 Go ahead and look for a good camping
 place for tonight
 How many boys deserted last night?*

Oi edeheto a?
 Ede bamona oi aena utua?
 Sibona ia gudu
 Toto oiemu vairana dekena idia muramura
 atoa?
 Lau kukuri neganai rara ia diho
 Lau bogana ia isisi
 Lau huahua negena kemaia isisi
 Oi mai iniseni
 Daekau mai iniseni
 Diko [=diho] lao unuseni
 Laegu hereva oi (umi) kamona
 Oi (umi) diba
 Edena negai ia ginidae?
 Lau boiboi negana oi (umi) mai haraga
 Ia revareva siai-a?
 Gabea lau revareva siai-a
 Badina dahaka memero idia heatu?
 Mero ta ia henaohenao
 Mero ta ia gau ta henao
 Memero idia lao labana
 Ina ... laohaia
 Sitoa dehena dekena atoa
 Una maua mailaia
 Kerukeru dabairere ita ai lao
 Nega hida ita (ai) mahuta ina dala
 dekena?
 Ororo badadia hida ita (ai) hanaia?
 Sinavai hida, ina dala dekena ia noho?
 Sinavai hida be badadia?
 Sedira haida ia abata?
 Oi lao guna hari hanuaboi mahuta gabu
 namona tahua
 Varani hanuaboi memero hida idia heau

*See that these things are not spoilt
by the rain*

*Have these packages (bundles, parcels)
prepared for carrying by the boys*

*Watch these things in case someone
steals them*

*The lashings are not secure and may
come undone*

*The men with the guns are coming
behind*

Are they far behind?

Erect the tents here

One boy cut some firewood

*One boy draw some water from the
creek*

Have the boys had their meal?

WORK GENERALLY

Have the boys gone to work?

Where are the boys working?

Tell them to cease that job

*The boys will cut grass this after-
noon*

*Five boys will dig this drain a
little deeper*

*The boys will come to the store at
midday*

What have you been doing today?

You go and get the axe

The axe has been lost

Go and search for the axe

I have searched in vain

How many knives are there?

Give me two knives

These knives are blunt

Have the knife sharpened

Whose knife is this?

Ina gauna naria namonamo, medu ia
hadikaia garina

Ina ikumi karia [=karaia] namonamo memero
idia huaia gwauraia

Ina gaudia naria, tau ta ia henaio garina

Ina waro idia kwatua namonamo lasi,
sibona ia ruhaia garina

Ipidi huaua taudia idia mai murinai

Murinai duadau idia mai?

Palai rumana karaia iniseni

Mero ta lahi taina siria

Mero ta diho sinavai ranu utua

Memero icia [=idia] aniani vadani

Memero idia lao gaukaraia?

Memero edesini idia gaukaraia?

Hamaoro-a idia hadokoa

Hari adorahai memero idia kurukuru (rei)
utua

Memero ima idia ina lesi (guri) geia
sisina dobu

Dina tubu memero idia mai sitoa dekena

Hari oi (umi) dahaka karaia?

Oi lao ira abia

Ira ia boio

Oi lao ira tahua

Lau tahua mose [=mase]

Kaia hida ia noho?

Keia rua oi henigu

Ina kaia matana lasi

Kaia matana segea

Ina daikena kaia?

Appendix 4: Excerpt from Nanai Gigovi's wartime propaganda recording

This excerpt is transcribed using standard Police Motu orthography. In addition the following symbols are used:

- ... speech continues but is interrupted by another
- faltering, false start, change of mind
- ,
- separates utterances that may be closely linked phonologically into grammatically meaningful units
- .
- indicates the end of sentence; also separates multi-word glosses for a single Police Motu word or morpheme

As well speakers are identified as follows:

- I: Interviewer
N: Nanai Gigovi

TAPE P1/83, SIDE 1:

- I: Umui lao negana dahaka umui itaia, tau kurokuro hanua dekenai?
you(pl.) go time what you(pl.) see man white village at
- N: O, gau momo idia karaia, tau kurokuro. Guna idia hereva, momo, diba
oh thing plenty they make man white before they say plenty know
- lasi taudia, idia gwau inai aniani gaudia ibounai, ipidi ibounai, wosipi
not persons they say this food things all gun all warship
- ibounai, stima ibounai, tau kurokuro ia karaia lasi, hari be mase taudia
all boat all man white he make not now focus dead people
- idia karaia. Lasi, mase taudia idia karaia lasi. Lau Nanai lau ura hereva,
they make no dead people they make not I Nanai I want say
- lau Nanai Gigovi, lau idiena hanua ibounai lau loaloa, lau itaia
I Nanai Gigovi I their village all I walk.around I see
- gau ibounai idia karaia. Misinigani idia karaia. Wosipi badadia idia
thing all they make machinegun they make warship big.ones they
- karaia. Plaimasisi ibounai idia karaia. Auri idia karaia. Tano bona auri
make aeroplane all they make iron they make soil and iron
- idia nadua, idia mistim. Ia daekau inai, laidi bamona. Vadaeni idia
they cook.up they mix.it it goes.up this light like okay they
- bubua tanika dekena. Idia abia lao misin dekena. Idia bubua, maraki maraki.
tip.it tank into they get go machine into they tip.it little little
- Vadaeni misini ia halataia ia lao bada herea. Ia gau latalata.
okay machine it makes.longer it go big very it thing longish
- Vadaeni idia kwatua.
okay they fasten
- I: Dahaka oi itaia tau kurokuro hanua dekena?
what you see man white village at
- N: Badina lau mai be Gavamani ia bogahisi maraki lasi bona Ostrelya
reason I come focus government it sad little not and Australia

taudia idia bogahisi maraki lasi lau dekena. Inai dekena lau mai,
people they sad little not I to this for I come
 idia-edia-ena hanua dekena.
their village to

FREE TRANSLATION

- I: *When you (pl.) went [to Australia] what did you see in the Europeans' towns [lit. villages]?*
- N: *Oh, they make lots of things, the Europeans. Before there used to be a lot of people who did not know what they were talking about who used to say that all this food and all these guns and warships and boats were not made by Europeans but were being made by [our] ancestors [lit. dead people]. No, [our] ancestors do not make them. I, Nanai, want to say that I, Nanai Gigovi, visited all their towns and saw all the things they make. They make machineguns, big warships, all the planes, iron. They cook up soil and iron and mix it up. This flares up like a light. Then they tip it into a tank. They take it in a machine and pour it out a little bit at a time. Then the machine extrudes it and it comes out much longer. Then they fasten it.*
- I: *What did you see in the Europeans' towns?*
- N: *The reason I went [lit. came] was because the Government and Australians were quite sorry for me. Consequently I went to see their towns.*

Appendix 5: Two examples of Nanai Gigovi's English

This speech is transcribed in a rough broad phonetic script based on English orthography except for *é* which is used to represent schwa. Otherwise ... is used to represent 'unintelligible sequence' and (?) = 'uncertain', N = Nanai and T = Tom [Dutton].

EXAMPLE 1 [Tape P1/84b: 207-225]

- | | |
|--|---|
| N: Tom, hau yu gedaun hia? Bai
plen o bot? | N: <i>Tom, how did you get down here? By
plane or boat?</i> |
| T: By plane. | |
| N: Wé frm? | N: <i>Where from?</i> |
| T: From Kerema. | |
| N: Orait. Yu kams t ^é miti mi a? | N: <i>All right. You came to meet me eh?</i> |
| T: Yes. | |
| N: Ana wi kem daun strets. | N: <i>And we came down straight away</i> |
| T: Nn. | |
| N: An yu luk mi fo sambadis an dei
kem daun. Ai hav a miting de fo
mai sirkets yunaitetsers. | N: <i>And you got somebody to look for me
and they came down. I was having a
circuit meeting there for the United
Church.</i> |
| T: United Church ... | |

N: Ana dei kam leti mi no au yo plens ede an ai gona mit yu de. Ana ai sed, "Orait." "Wea goin daun t^é yu houm, Babaguna," an ai tel yu ai gona sida mai dota bikos i hava mai mota. Nau ai gona sida mai dota se sed, "O gasget iz bloaut. Motas iz no gut, iz nati wokin. Mai asbini noti beiki yet andes stil ina fishin." Nau we kems an yu sed, "At wan klok we goin daun an gamant i help." Mota we bring ^ét ana yu shop an yu pei da ah -- benzin an wi kem t^é, tu galen, wi kem daun hia ana draiva lutin (?) an i riten tu moning at eit klok ah -- nein klok bi hia. An wi teika yu gen. An nau wi ala go bek intu an Kikori an yu gouen tu Baimuru ana yu alredi liv mi ona Kikori.

T: Very good. Where di you learn English?

N: O ai len hia -- wen ai ovasia. Ai workin long long taims wen aimz a litl boi. Aiz a kuk, Gamant -- kuk, hausbois.

T: In town?

N: No, no, no. Tauns not. Desa Kikori de, gavman tesin. O bifo ... kantri ful ... patrol offisas, an eidisi an ovasis. Nau wi Papua Niugini kam t^é reples em Kikori ti(?) patrol ofsas.

T: Do you remember Mr Harris?

N: Yea, ai no. He kem an stei de with mai brada. He kem state -- wotyukol^ét -- dei shut krokadail -- araun de ai with ah -- Eipisi -- ai dzes hed abt wan yuropin daun t^é yo ples ai kem ... dei stei de meki houm ... neks taim.

T: He was a patrol officer here.

N: No, noti patrol ofisa ... he dzes kam tu shuti krokadails an ...

N: *And they came and let me know that your plane was there and I'd be going with you. And I said, "All right." "We're going down to your place to Babaguna." And I told you that I'd go and see my daughter because she had my [outboard] motor. Then I went and saw my daughter and she said, "Oh the gasket has blown out. The motor's no good, it's not working. My husband is not back yet. He's still out fishing." Then we came and you said, "At one o'clock we're going down and the Government will help." And we brought the motor [down] and you shopped/ stopped (?) [at the store] and paid for the ah-petrol and we came -- two gallons -- we came down here and the driver returned and he's to return again in the morning at 8 o'clock ah -- 9 o'clock to be here. And we'll pick you up again. And then we'll go back into Kikori and you'll go to Baimuru and you will have left me in Kikori.*

N: *Oh, I learnt here when I was an over-seer. I was working for a long time when I was a little boy. I was a cook, Government -- cook, houseboy.*

N: *No, no, no. Not in town. Just in Kikori there, at the government station. Oh, before I had been with patrol officers, ADC [Assistant District Commissioner] and [people] from over-seas. Now we Papua New Guineans have come to replace patrol officers at Kikori.*

N: *Yes, I know. He came to stay there with my brother. They shot crocodiles. I was with APC. I just heard about a European being down there at my place and I came ... they stayed there and made their home ... next time.*

N: *No, not a patrol officer. He just came to shoot crocodiles and ...*

EXAMPLE 2 [Tape P2/82a: between 100 and 200]

Nanai is here recalling what happened to him when he first arrived at Port Moresby during the war. He was met at the wharf by an army officer [=A] who said:

- | | |
|-------------------------------------|---|
| A: Nanai de? | A: <i>Is Nanai there?</i> |
| N: Yea, am hia. | N: <i>Yes, I'm here.</i> |
| A: Orait, kamon gedon. | A: <i>Okay then, come and get on.</i> |
| N: O lau be lau gari. | N: <i>Oh, I'm frightened.</i> |
| A: Dasmata. | A: <i>That doesn't matter.</i> |
| N: Lauegu suitkes lau rakatania. | N: <i>Will I leave my suitcase behind?</i> |
| A: No wari, liv it. Oi mai. | A: <i>Don't worry. Leave it. Come on.</i> |
| N: Ah, edenabamona? Wadan stretiap. | N: <i>Heh, how? Okay [we went] straight up.</i> |

Appendix 6: List of words used for survey purposes

The following list was compiled mainly from the wartime written materials discussed in this paper. It is not exhaustive (that is, does not include every item considered questionable by me for some reason) but was used as a base to obtain reactions and other information.

- | | |
|---|---|
| <i>fruit</i> (anani) | <i>bow (weapon)</i> (pau) |
| <i>mountain</i> (aulolo) | <i>ground</i> (tanu) |
| <i>hill</i> (otolo) | <i>boiling water</i> (wa siau) |
| <i>call s.o.</i> (boidia) | <i>mirror</i> (wariwari) |
| <i>bottle</i> (botol) | <i>sneeze</i> (asimana) |
| <i>hold something</i> (dogwata) | <i>prostitute</i> (ariara hahine) |
| <i>if</i> (ena gena) | <i>we(excl.)</i> (ai) |
| <i>shark</i> (gabuli) | <i>guard</i> (gadi tauna) |
| <i>turtle, tortoise</i> (kapori) | <i>cookboy</i> (kuki tauna) |
| <i>reef</i> (gadu bada) | <i>sago needle</i> (kinkin) |
| <i>report</i> (hamaru) | <i>parting with friend</i> (ahetoni) |
| <i>their</i> (idiena) | <i>angel</i> (aneru) |
| <i>our(incl.)</i> (itena) | <i>flower of tree</i> (au burena) |
| <i>always</i> (hanahanai) | <i>to weed</i> (ava raua) |
| <i>trochus shell</i> (ilama, toea, mairi, bisisi) | <i>lungs</i> (baragi) |
| <i>stupid, silly</i> (kava masi) | <i>gardening</i> (biru) |
| <i>lock up</i> (kikaraia) | <i>to snatch</i> (biru) |
| <i>biscuit</i> (kiskis) | <i>corner</i> (daiguni) |
| <i>chook, fawl</i> (kokoroga, kamkam) | <i>room</i> (daiutu) |
| <i>fish</i> (kwaremi) | <i>coconut oil</i> (dehoro) |
| <i>repeat</i> (kworaia lo) | <i>to lick</i> (demaria) |
| <i>sipoma</i> (levu) | <i>temptation</i> (dibagani) |
| <i>rubber</i> (magi) | <i>how</i> (edeheto) |
| <i>sleepy</i> (mahuta masi) | <i>barren</i> (gabani) |
| <i>mosquito</i> (namu) | <i>coition</i> (gagaia, lagaia) |
| <i>beg</i> (noienioe) | <i>to watch</i> (gima) |
| <i>milk</i> (rata, susu) | <i>promise</i> (gwahamata) |
| <i>tree fern</i> (otolo)(hatoro, vatolo (Rigo)) | <i>answer</i> (haere henia, ansa henia) |
| | <i>glory, pretty</i> (hairaina) |
| | <i>boast</i> (heagi) |

to race (heau helulu)
twins (hekapa)
to covet (hekisehekisi)
a paddle (hode)
adultery (heudahanai)
all of us (iboudai)
bitter (idita)
saw (iri)
tears (irurumata)
a fan (itapo)
pinch (kinia)
to hide (komu)
yaws (kuhi)

breath (lababana)
fence (magu)
to drown (maloa)
greedy (matagangani)
dumb (mu)
to wait in vain (naria mase)
double canoe (puapua)
pull up (ragaia)
to court (sihari)
umbrella (tamaru)
just so, all right, therefore (taunabinai)
to mock, copy (tohotoho)
dusk, dark (vabua)

Appendix 7: Some common English-based words in Police Motu

The following list contains common English-based words that form a significant part of the vocabulary of Police Motu but which are generally overlooked in stressing the Motu nature of Police Motu. In many cases there are common Motu alternatives available which are used in the Port Moresby area if they are known or the item can be described periphrastically in Police Motu if one wants to avoid these, e.g. kabis *cabbage, greens, vegetables* can be described as uma gabu dekena hadoa aniani (lit. *garden place in planted food*). However, Police Motu speakers outside of Port Moresby speaking amongst themselves will generally use these English-based forms in preference to such Motu-based ones. Most of those to be listed are pre-war in origin but there are many more that are post-war in origin. The list does not claim to be exhaustive.

ENGLISH-BASED FORM	COMMON MOTU-BASED FORM	MEANING
aiglas	(mata)galasi	<i>spectacles</i>
aian		<i>(clothes) iron</i>
ais, aisi		<i>ice</i>
amerika tauna		<i>an American</i>
ami	tuari taudia	<i>army</i>
anka	dogo	<i>anchor</i>
ansa henia	haere henia	<i>to answer</i>
apolo, apul		<i>apple</i>
au plaua plaua	au burena	<i>flower</i>
aut	doko, haorea	<i>out, finished</i>
autbod (motor/vanagi)		<i>outboard (motor/canoe)</i>
autsait	gadobada	<i>open sea beyond the reef or bar</i>
aven	amudo	<i>oven</i>
Baibel		<i>Bible</i>
baketi		<i>bucket</i>
bampa	kerere, tatakau, kamokau	<i>bump, run into, have an accident</i>
bandesi		<i>bandage</i>
baraki	polisi taudia edia ruma	<i>barracks, rest house</i>
baranda	dehe	<i>verandah</i>

ENGLISH-BASED FORM	COMMON MOTU-BASED FORM	MEANING
batari		<i>battery</i>
beig, bek	puse	<i>bag, sack</i>
bekhaus	plaua karaia rumana	<i>bakery</i>
bensin		<i>benzine, petrol</i>
beredi	plaua	<i>bread</i>
bia	muramura	<i>beer</i>
bilanket(i)	goru, geda	<i>blanket</i>
Bipi		<i>B.P. (Burns Philp Pty.Ltd.)</i>
biugli	kibi	<i>bugle, trumpet</i>
bloaut	poua	<i>blow out, break through, explosive force, burst</i>
Britani		<i>Britain</i>
boila	uro (badana ta)	<i>boiler</i>
bolo		<i>ball</i>
bomu [wartime introduction]		<i>bomb</i>
bosi tauna, bosboi, poman	biaguna	<i>boss, overseer, foreman</i>
busikasi, pusikasi		<i>cat</i>
botol	kavabu	<i>bottle</i>
botm ia abia	dogo	<i>to hold fast in a boat</i>
bouti	vanagi	<i>boat</i>
brum	daroa	<i>to sweep</i>
buka		<i>book</i>
bulamakau, boromakau	vamu	<i>meat, beef, cattle</i>
daivinglas		<i>diving mask/goggles</i>
dasta	kahu	<i>dust</i>
dablvanagi, lakatoi	puapua	<i>double-hulled canoe</i>
daki, dakdak		<i>duck</i>
dingi	vanagi	<i>dinghy</i>
dioni, sioni	hakupua	<i>to join (timber)</i>
distrik	kahana, gabuna	<i>district, area</i>
dokta tauna	muramura henia tauna	<i>doctor</i>
draiva tauna	taria tauna	<i>driver</i>
elemes		<i>LMS (London Missionary Society)</i>
ensini, insini		<i>engine</i>
estrip	plein ia diho/helai gabuna	<i>airstrip</i>
fadom		<i>fathom (in water)</i>
faiv	ima	<i>five</i>
fakteri	kohu karaia gabuna	<i>factory</i>
famili	bese	<i>family</i>
fes, nambawan	ginigunana	<i>first</i>
faia	poua, hapoua	<i>to fire (a gun)</i>
fens	ara, magu	<i>fence</i>
flot	hure	<i>to float</i>
foka	du gadana, au gadagadana	<i>fork, house post</i>
foto		<i>photograph</i>
fraipan		<i>frying pan</i>
frisa		<i>refrigerator</i>
gadi tauna	gima tauna	<i>guard</i>
gavamani, gamanti		<i>government</i>
gavana		<i>governor, administrator</i>
gita		<i>guitar</i>
golo		<i>gold</i>

ENGLISH-BASED FORM	COMMON MOTU-BASED FORM	MEANING
glas, galasi		<i>glass, any glass used for seeing, diving, measuring, drinking</i>
got, nanigos/t		<i>goat</i>
halusia		<i>waste, squander</i>
hambaga tauna		<i>a nuisance person</i>
hanwas	dina gauna	<i>handwatch</i>
hapu	kahana	<i>half, piece of</i>
hawa	hora	<i>hour</i>
hosi		<i>horse</i>
inglan		<i>England</i>
inglisi		<i>English</i>
intepira - v. toropeta		
ius(laia)	abia	<i>to use</i>
kabis	uma gabu dekena hadoa aniani	<i>cabbage, greens, vegetables</i>
kada	vanagi	<i>cutter, boat</i>
kala		<i>colour</i>
kalenda		<i>calendar</i>
kapenta	ruma karaia tauna	<i>carpenter</i>
kapten	biaguna	<i>captain, leader</i>
kapusi		<i>cup</i>
kas(i)		<i>card</i>
katris, katresi, kasiresi		<i>cartridge, bullet</i>
kaunsila		<i>councillor</i>
katolik		<i>Catholic</i>
keke		<i>cake, biscuit</i>
kempa	hanua maraki ta	<i>camp</i>
kerosin		<i>kerosine</i>
kini	pavapavana	<i>king</i>
kiskis, biskes/ti		<i>biscuit</i>
koki	karai	<i>cockatoo</i>
kopi		<i>coffee</i>
kota karaia		<i>to take to court</i>
kuki mero/tauna	nadua mero/tauna	<i>a cook (boy)</i>
kwini	pavapavana	<i>queen</i>
lain karaia		<i>to line up, parade</i>
laisens		<i>licence</i>
laki		<i>lucky, name of card game</i>
lampa	lahi	<i>lantern, lamp</i>
leba taudia	gaukara taudia	<i>labourers</i>
leiti	mai murinai	<i>late</i>
leta	revareva	<i>letter</i>
livi		<i>leave, vacation</i>
liva	asena	<i>liver</i>
lodi badana		<i>a big load</i>
lori		<i>truck, lorry</i>
lukinglas	wariwari	<i>mirror</i>
mande (and other days of the week)		<i>Monday</i>
manuwa, wosipi	tuari vanagi/lagatoi	<i>warship</i>
masisi	lahi	<i>matches</i>
medikolo tauna		<i>medical orderly</i>

ENGLISH-BASED FORM	COMMON MOTU-BASED FORM	MEANING
mei (and other months of the year)		<i>May</i>
minin	anina	<i>meaning</i>
misin		<i>mission, machine</i>
mistim	giroagiroya	<i>mix</i>
moni	davana	<i>money</i>
moto		<i>motor</i>
namba		<i>number</i>
nambawan, fes	ginigunana	<i>first</i>
nambatu, seken	iharuana	<i>second</i>
nanigos - v. gos		
nes		<i>nurse</i>
nila	ginigini	<i>needle</i>
nius	hereva	<i>news</i>
niuspeper		<i>newspaper</i>
oda		<i>order (goods), give an order</i>
oila		<i>oil</i>
opis		<i>office</i>
opisa		<i>officer</i>
ostrelya tauna		<i>an Australian</i>
paipu	baubau	<i>pipe</i>
pait, paizi	heatu, heai, tuari, botaia imana dekena	<i>to fight (with fists)</i>
pamu		<i>pump</i>
panisi		<i>punish</i>
papala		<i>propellor</i>
paua	goada	<i>power, electricity</i>
paun		<i>pound</i>
paura	kahu	<i>powder</i>
pegi, feg, feki		<i>peg</i>
pei	davana	<i>pay, wages</i>
pernit		<i>permit</i>
pepa		<i>paper</i>
penia	umua	<i>to paint</i>
pinat	niuniu	<i>peanut, ground nut</i>
pinin gado		<i>a pidgin language</i>
pisu	kahana, duduna, sisina	<i>a piece of</i>
plai, palai	lara	<i>a sail</i>
plaimasini, eloplen, plen, plein		<i>aeroplane</i>
plaua, palaua, beredi		<i>bread</i>
plein - v. plaimasini		
ploa	reirei	<i>floor</i>
poinia	gau ta duduia	<i>to point to</i>
pokapoka	du	<i>house post</i>
polisimani		<i>policeman</i>
polisel	mairi	<i>pearl shell</i>
poman - v. bositauna		
poren		<i>to fall in, parade</i>
pusikasi - v. busikasi		
raba	magi	<i>rubber, any instrument using rubber as a motive force</i>

ENGLISH-BASED FORM	COMMON MOTU-BASED FORM	MEANING
rabagan		<i>speargun</i>
raisi		<i>rice</i>
redi	naria, hagoevaia	<i>ready, prepared</i>
rin(i)		<i>a ring, to ring</i>
rip	moemoe, didi	<i>reef</i>
riva	sinavai	<i>river, stream</i>
rup	guhi	<i>roof</i>
saini		<i>a sign, to sign (a paper), to sign on</i>
saini karaia		<i>to make a sign</i>
sait(sait)	kahana(kahana)	<i>the side of</i>
sak(i)	kwalaha	<i>shark</i>
satap!	hereva lasi	<i>Shut up!</i>
sekap		<i>to check up</i>
seken - v. nambatu		
sel, sels	koukouna	<i>shell</i>
sens	haidaua	<i>to change</i>
senta momokani!	gauna inai!	<i>Spot on! That's it!</i>
seven	hitu	<i>seven</i>
sevendei tauna		<i>a Seventh Day Adventist adherent</i>
sevi	huihuina utua	<i>to shave</i>
sikis	tauratoi	<i>six</i>
simenti		<i>cement</i>
singlis		<i>singlet</i>
siok(a)		<i>chalk</i>
sipunu	bedi	<i>spoon</i>
sipeli lasi	ore lasi, doko lasi	<i>to continue without a rest</i>
sipi	mamoe	<i>sheep</i>
sioti, sieti		<i>shirt</i>
sisel		<i>chisel</i>
sisima	vanagi	<i>ship, steamer</i>
skel		<i>scales, to weigh out</i>
skul, sikulu	hadibaia ruma, hadibaia gaukara	<i>school (house), to go to school, to learn</i>
slin	vilipopo	<i>sling, catapult</i>
smokhaus		<i>smokehouse</i>
smolhaus	kukuri rumana	<i>toilet</i>
sno	ninoa, ori ia diho	<i>fog, snow</i>
so	iri	<i>a saw (instrument)</i>
soka	bolo gadara	<i>soccer</i>
sol	damena	<i>salt</i>
solpis		<i>salted fish</i>
somil	au utua gabuna	<i>sawmill</i>
soppis		<i>beche-de-mer, sea cucumber/slug</i>
sopu		<i>soap</i>
sori	madi	<i>Oh sorry!</i>
spidi momo	heau haraga	<i>to speed</i>
stanibai	naria	<i>be prepared, ready</i>
staka		<i>a lot of, a stack</i>
stakim	haboua	<i>to heap up</i>
stat	(ha)matamaia	<i>to begin, start</i>
stesin		<i>station, government post</i>

ENGLISH-BASED FORM	COMMON MOTU-BASED FORM	MEANING
stiki	au duduna	<i>a stick</i>
stoa, stua	hoihoi gabuna/rumana	<i>a store</i>
suga	tohu	<i>sugar</i>
supu	vasiahu	<i>soup</i>
susu	rata	<i>milk, latex</i>
sutkes, sutkeis	maua	<i>suitcase, box</i>
swampa	gohu	<i>swamp</i>
suit momo		<i>very sweet</i>
taipia		<i>to type</i>
taipraita		<i>typewriter</i>
tanika		<i>tank</i>
tauni	hanua	<i>town</i>
tenkiu	namo	<i>thank you</i>
tin(i)		<i>tin</i>
tisa (tauna)	hadibaia (tauna)	<i>teacher</i>
toea	ilama	<i>trochus shell</i>
taraka, lori		<i>truck, lorry</i>
trakta		<i>tractor</i>
toropeta, interpa	gado hanaia tauna	<i>interpreter</i>
tosu		<i>torch</i>
traim	verb + itaia	<i>to try</i>
trosel	ilama, toea	<i>trochus shell</i>
waia		<i>wire</i>
wailis		<i>wireless</i>
wil		<i>wheel</i>
windo	gaba mauru	<i>window</i>
wiski	muramura	<i>whisky</i>
wol	haba	<i>wall</i>
wolbot	vanagi	<i>whaleboat</i>
yelo	laboralabora	<i>yellow</i>
yia(s)	lagani	<i>year(s)</i>

Appendix 8: Some unusual non-English-based words in Police Motu

The following list contains words that are unusual because: a) they are derived from languages of Papua other than Motu; b) they are different in form from those normally found in vocabularies or dictionaries of Police Motu. Some of these are areally distributed. Where this is the case the relevant area is shown in the column on the righthand side and the following abbreviations are used:

- E: Eastern Papua
- W: Western Papua
- C: Central Papua

Where no area is shown the item is regarded as being widely distributed.

ITEM	SOURCE LANGUAGE	MEANING	FORM NORMALLY GIVEN IN VOCABULARIES OR DICTIONARIES OF POLICE MOTU	AREA
bara	Motu	<i>a paddle</i>	hode	
boiboi	Motu	<i>to call someone</i>	boiria	
diho ranu/ davara dekana	Motu	<i>to drown</i>	maloa	
dimdim	Suau(?)	<i>white man, European</i>	tau kurokuro	E
edena bamona	Motu	<i>how</i>	edeheto	
gauna inai	Motu	<i>that's it, all right</i>	taunabunai	
gebore	Motu	<i>earring</i>	taia gauna	
geda	Motu	<i>rain cape, umbrella (in areas where mats (geda) are so used)</i>	damaru	
goru	Motu	<i>umbrella, rain cape (in areas where the leaf sheath of the goru palm is so used)</i>	damaru	
mamoos, mamus	Torres Straits	<i>chief, village policeman</i>	hanua lohia, hanua polisiman	W
hereva maraki, maraki	Motu	<i>to whisper</i>	gado maraki, hereva metaira metaira	
idiena	Motu	<i>their</i>	idia-edia	
iduaru	Motu	<i>gate</i>	ikoukou	
iena	Motu	<i>his, her, its</i>	ia-ena	
kamkam	Suau(?)	<i>fowl, bush fowl</i>	kepoka	E,W
kapani	Motu	<i>woman whose husband has left her</i>	hadihoa	
kapore	Suau(?)	<i>oh sorry</i>	o madi	E,W
katoro	?	<i>pandanus nut</i>	geregere	C inland (Goilala)
kavakava tauna	Motu	<i>dumb</i>	mu	
kolohau	Motu	<i>shark</i>	kwalaha	
koreme	Motu	<i>fish</i>	gwarume	
kukumara	?	<i>exchanging tobacco to settle dispute</i>	?	C inland (Goilala)
kurokuro dikana	Motu	<i>weeds</i>	ava	
kurokuro dikana kokia	Motu	<i>to weed</i>	ava raua	
lakatoi	Motu	<i>double-hulled canoe</i>	puapua	
matana ranu	Motu	<i>tears</i>	irurumata	
mutu, diho ranu dekena	Motu	<i>to drown</i>	maloa	
natuna lasi hahine	Motu	<i>barren (of woman)</i>	gabani	
natuna rua	Motu	<i>twins</i>	hekaapa	
nemo	Motu	<i>mosquito</i>	namo	
omi, umi	Motu	<i>you (pl.)</i>	umui	
ruma maraki	Motu	<i>a room</i>	daiutu	
siari (kava) kekeni/hahine	Motu	<i>prostitute</i>	ariara hahine	
susu	Suau(?)	<i>milk, juice</i>	rata	
udubaubau toto	Motu	<i>yaws</i>	kuhi	
uma gabu gaukara	Motu	<i>gardening</i>	biru	

REFERENCES

BARRETT, Don

- 1969 The Pacific Islands Regiment. In K. Inglis, ed. *The history of Melanesia*, 493-502. Canberra: The University of Papua New Guinea in association with the Research School of Pacific Studies, The Australian National University.

BRETT, Richard, Raymond BROWN, Ruth BROWN and Velma FOREMAN

- 1962 *A survey of Motu and Police Motu*. Port Moresby: Department of Information and Extension Services.

CHATTERTON, Percy

- 1946 *A primer of Police Motu*. Cairns: The Cairns Post.
- 1950 *A primer of Police Motu*. Port Moresby: The Department of Education.
- n.d. [ca. 1959] *A primer of Police Motu*. Sydney: Pacific Publications. Reprinted ca. 1968.
- 1971 *Hiri Motu (Police Motu)*. Port Moresby: U.F.M. Press.
- 1975 *Say it in Motu: an introduction to the common language of Papua*. Sydney: Pacific Publications.

DUTTON, T.E.

- 1969 *The peopling of Central Papua: some preliminary observations*. PL, B-9.
- 1976 Review article: *The dictionary and grammar of Hiri Motu* (Port Moresby: The Office of Information), *Kivung* 9/1:39-56.
- 1983 Birds of a feather: a pair of rare pidgins from the Gulf of Papua. In E. Woolford and W. Washabaugh, eds *The social context of creolization*, 77-105. Ann Arbor: Karoma.
- 1985 *Police Motu: iena sivarai*. Port Moresby: UPNG Press.

DUTTON, T.E. and C.L. VOORHOEVE

- 1974 *Beginning Hiri Motu*. PL, D-24.

LISTER-TURNER, R. and J.B. CLARK

- n.d.a *A grammar of the Motu language of Papua*. Second edition, edited by Percy Chatterton. Sydney: Government Printer.
- n.d.b *A dictionary of the Motu language of Papua*. Second edition, edited by Percy Chatterton. Sydney: Government Printer.

LONG, Gavin

- 1973 *The six years war: a concise history of Australia in the 1939-45 war*. Canberra: The Australian War Memorial and the Australian Government Publishing Service.

McCARTHY, Dudley

- 1959 *South-West Pacific area - first year: Kokoda to Wau*. Canberra: The Australian War Memorial.

MÜHLHÄUSLER, P.

- 1978 Papuan Pidgin English rediscovered. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings, 1377-1446.* PL, C-61.

MURRAY, J.H.P.

- 1930 Foreword. In R. Lister-Turner and J.B. Clark *Revised Motu grammar and vocabulary*, pp.iii-v. Port Moresby: Government Printer.

NELSON, Hank

- 1982 Taim bilong pait na taim bilong hatwok tru: Papua New Guinea labourers during the war, 1942-1945. MS, Department of South-East Asian and Pacific History, Research School of Pacific Studies, The Australian National University.

SHNUKAL, Anna

- 1983 Torres Strait Creole: the growth of a new Torres Strait language. *Aboriginal History* 7:173-185.

WURM, S.A.

- 1975 The Trans-Fly (Sub-Phylum Level) Stock. In S.A. Wurm, ed. *New Guinea area languages and language study, vol.1: Papuan languages and the New Guinea linguistic scene*, 323-344. PL, C-38.

REFUGEES, TRADERS, AND OTHER WANDERERS:
THE LINGUISTIC EFFECTS OF POPULATION MIXING IN MELANESIA

Ann Chowning

0. INTRODUCTION

In the conclusion to his famous paper on indirect inheritance in Rotuman, Biggs made the following points:

In general what we know of culture history in the Melanesian area suggests a complex rather than a simple linguistic history, involving a good deal of movement in certain maritime areas, and long continued contact among speakers of related languages ... and unrelated languages It would be surprising indeed if such contacts did not have substantial effects upon the languages concerned, effects which could be vitally important to comparative work, and to our understanding of Pacific prehistory. While talk of substrata and mixed languages may in fact introduce concepts which are both ill-defined and unhelpful, the multiple origins of Melanesian lexicons, if real, should be studied. Failure to examine the extent to which one language has been affected by others can lead to erroneous subgrouping.

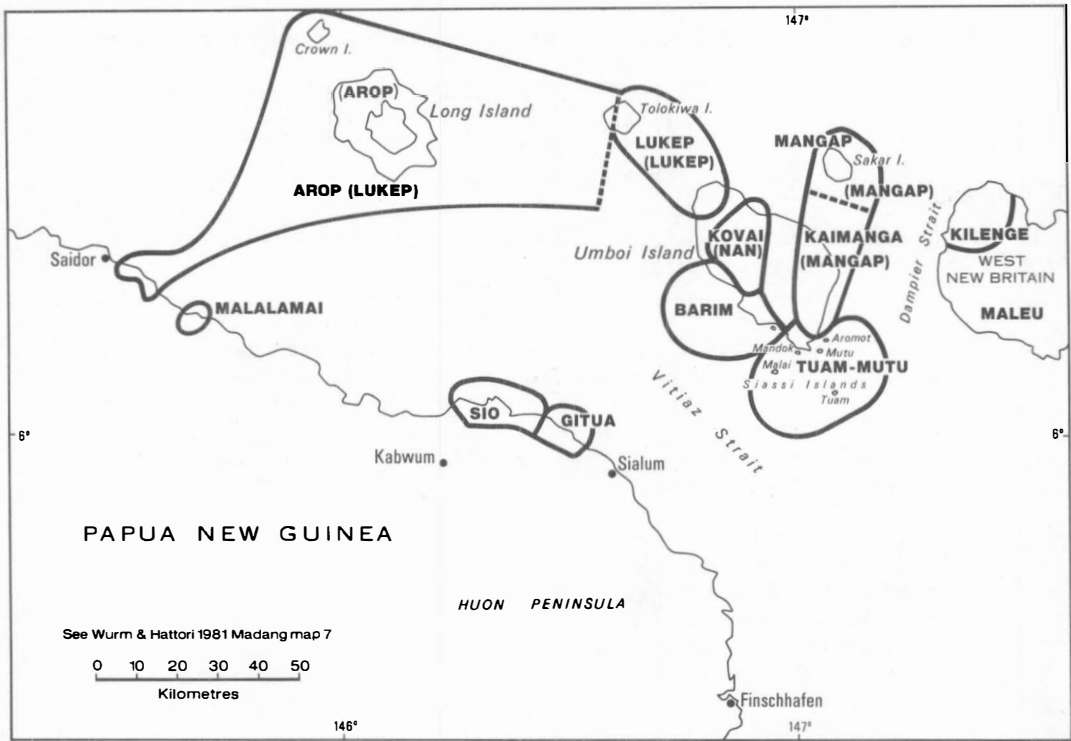
(Biggs 1965:414-415)

Yet little attention has been paid to these warnings. Those interested in subgrouping Melanesian languages, myself included, have tended to assume that languages sharing a substantial number of obvious cognates belong to the same subgroup as descendants of a single ancestor. The inevitable reliance on short wordlists often prevents the establishing of regular sound correspondences, and apparent irregularities may be ignored on the assumption that most could be explained if more data were available. Sometimes, however, additional data actually complicate the picture, raising questions about the nature and identity of the putative common ancestor. Such complications emerged when I began to examine in detail certain languages of the region between north-west New Britain and the north coast of New Guinea. This general region is renowned for its linguistic diversity (Dyen 1965), but there is still much disagreement about the numbers and boundaries of the subgroups (see Hooley 1976:341-344; Z'graggen 1976: 287; Ross 1977; Lincoln 1977b). Although some of the disagreements reflect different criteria for subgrouping, others derive from the linguistic situation itself. The possible reasons for the situation found around the Vitiav Strait will also be offered as helping explain difficulties found in subgrouping certain

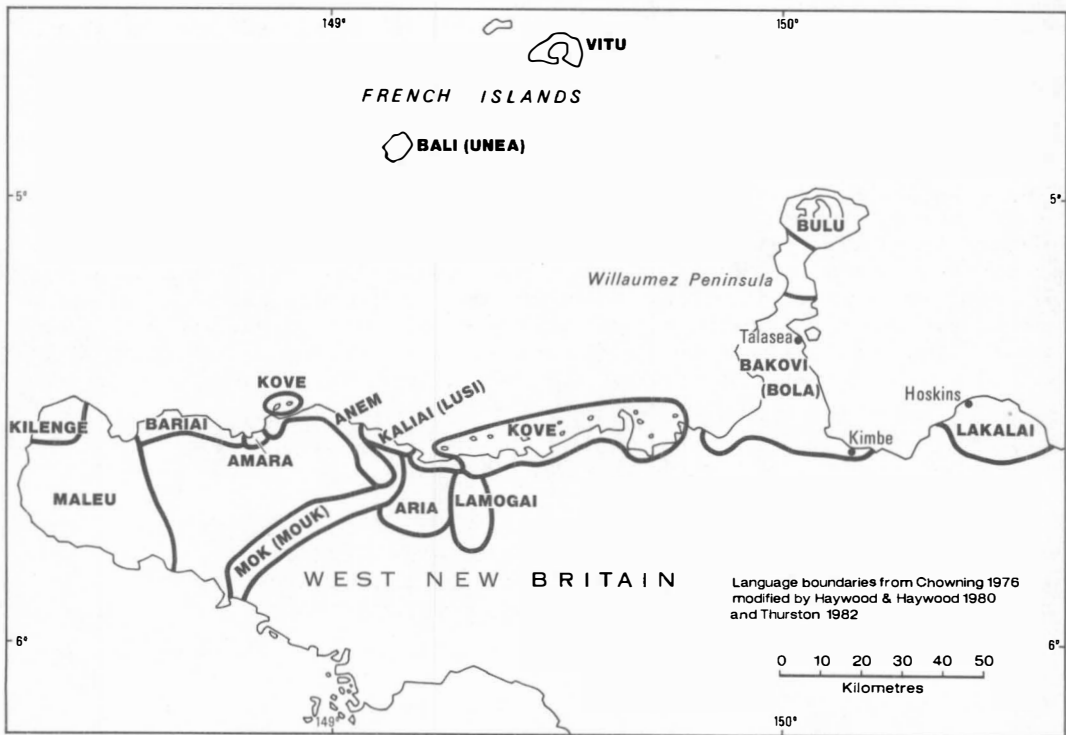
Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 407-434. *Pacific Linguistics*, C-94, 1986.

© Ann Chowning

407



Map 1



Map 2

languages found farther east, such as those of the French Islands and of parts of Milne Bay.

The discussion will begin with a description of postulated linguistic ties between New Britain and New Guinea, followed by a more detailed analysis of the resemblances and differences between two seemingly close relatives, Gitua of New Guinea and Kove of New Britain. These will then be compared with other languages in the region, in an attempt to understand why their relationship is probably not so simple as it once seemed.

1. KOVE AND GITUA

In this section the question of the nature of the relationship between these two languages will be investigated.

1.1 Background

In 1971, Hooley suggested that almost all the coastal AN languages spoken in what is now the Morobe Province belonged to a single family, which he called Siasi,¹ and that this had several subfamilies. The largest of these he labelled the Island Subfamily because most of the languages are located on offshore islands, including those of the Siassi group between New Guinea and New Britain. Gitua was put into this subfamily, while Maleu off the western tip of New Britain was assigned to the Siasi Family but not to the Island Subfamily. The Siasi Family was also stated to extend westward into Madang Province (Hooley 1971:99-104). In 1973 I argued that Maleu was a somewhat aberrant member of what I had been calling the Bariai Family of north-west New Britain, comprising most of the AN languages (Kove-Kaliai, Bariai, and Kilenge-Maleu) spoken along the coast from the western side of the Willaumez Peninsula to Cape Gloucester. I suggested that Kove, and therefore Maleu as well, clearly belonged to the same subgroup as Gitua, Tuam-Mutu of the Siassi Islands, and probably Malalamai of the Madang coast (Chowning 1973:208-209). Hooley accepted my argument (1976:344). In 1976, however, Lincoln undertook fieldwork on the Morobe coast, combining a lengthy stay in Gitua with a survey of all the coastal AN languages which Hooley had assigned to the Siasi Family. Lincoln agreed with me that Kove belonged in the same subgroup with Gitua and Malalamai, but argued for a separation of Maleu-Kilenge from the other Bariai languages, as well as proposing many alterations to Hooley's subgroups. He suggested that Bariai, Kove-Kaliai, Gitua, and Malalamai belonged to one subgroup, which he tentatively named *Ŋero*, while various other languages of the small islands, along with Kilenge-Maleu, belonged to a different "Sio" subgroup (Lincoln 1976a, 1977b). At this point, the only person questioning the close relationship between Kove and Gitua was Ross (1977); he "very tentatively" suggested on purely phonological grounds (some mistaken) that Kove should be separated from the *Ŋero* languages. His suggestions will be discussed later.

Meanwhile, Lincoln had supplied me with a considerable amount of Gitua material, including much lexical and grammatical data (Lincoln 1977a, 1977c), and I had become increasingly uneasy about the many irregular sound correspondences between the two languages. There was no doubt the Kove and Gitua were linked by many isoglosses, some involving items that were very uncommon in this region. The difficulty lay in constructing proto-forms from which the shared

items might be derived. Before discussing the problems, examples of the shared items will be given. The items noted here either seem not to derive from POC forms or show unexpected phonological innovations.²

1.2 Shared lexical items

Rather surprisingly, in view of the trade network that links the Vitiaz Straits region (Harding 1967), and the distribution of a common art style (Bodrogi 1961, Dark 1974), the shared forms include very few cultural items. Furthermore, they also include very few names of animals and plants. A partial explanation lies in the very different flora and fauna of New Guinea and New Britain (not to mention the reduced biota of the Siassi Islands), but the rarity of shared names for sea creatures is less explicable. Instead, the list contains many shared names of everyday objects, including parts of the body; verbs; numerals; and negatives. Sound correspondences will be discussed below. Obvious cognates include the following: Gitua *bwae*, Kove *voe armpit*; B. *bwale fleshy part of buttocks*, K. *vole hip*; G. *guvi*, K. *yuvi heel*; G. *tuzu*, K. *туру breast*; G. *apwa*, K. *apo belly*; G., K. *ŋoŋo nasal mucus*. Most of the numerals are derived from POC. The most distinctive one is G., K. *paŋe four*; see also G. *eze*, K. *ere one*. Both negatives are cognate (G. *mago*, K. *mao no*; G., K. *mina don't*), and both follow the words they modify. Examples of many other cognates will appear below.

The two languages also share what seems to be a rare morphophonemic alternation of initial consonants in some of the same pairs: G. *poze*, K. *pore paddle* (n.); G. *voze*, K. *ore to paddle*; G. *sage*, K. *sae above*; G. *zage*, K. *rae go up* (see Chowning 1973:200 for the list of such alternations in Kove). In one case, Gitua seems to show an alternation which Kove lacks: Gitua *van give*, *pan to*, *from*, etc. Kove has *pani* for both of these.

1.3 Sound correspondences

Kove phonemes were described in Chowning 1973:194-195, but because Lincoln and Ross have emphasised what they see as Kove peculiarities, some points need repeating. Kove has a set of voiced fricatives, written as *ɣ*, *r*, and *v* (respectively velar, alveolar, and bilabial), which usually correspond to voiced stops in Bariai (see further discussion under 3.2 and 4.). The /r/ is often pronounced as an alveolar trill in the western dialect of Kove,³ spoken adjacent to Kaliai, which has two r-phonemes, one a flap and one a trill. The latter usually corresponds to /h/ in Kove. The trilled pronunciation may represent Kaliai influence. The alveolar fricative deserves some attention because it often corresponds to an alveolar fricative in Gitua which Lincoln has written as *z*. It is not known how much these two phonemes actually differ in pronunciation. The question of the correspondences will be discussed further below.

In Kove and Kaliai, as in Gitua, the velar fricative may be pronounced as a voiced stop in word-initial position (see Counts 1969:18). Probably for this reason, Lincoln has chosen to write the voiced velar stop which is a separate phoneme as a digraph, *ŋg*. The Kove voiced stops are also prenasalised word medially. It seems that Kove /ɣ/ is pronounced like Gitua /g/, and Kove /g/ like Gitua /ŋg/.

Finally, to clarify the material in Table 1, it should be noted that Lincoln has expressed (1976d) uncertainty about "the status of /w/ and /y/ as

separate phonemes". In Kove, /w/ certainly exists as a separate phoneme, but cannot usually be distinguished from pre-vocalic /u/,⁴ so that there is uncertainty about the spelling of some words. Kove /y/ may not be a separate phoneme, usually being heard as /e/ or /i/ in slow speech.

A notable difference between the languages is that many Gitua words end in a consonant, whereas word-final consonants are so rare in Kove that they almost surely occur only in recent borrowings. The only word-final consonants in Gitua are /k/, /p/, /m/, /n/, /ŋ/, /l/, and /r/. Often this word-final consonant reflects the loss of a final POC vowel, as in G. aŋar (K. aŋahi, POC *kaŋaRi) *canarium almond*; G. gan (K. ani, POC *kani) *eat*; G. tun (K. tunu, POC *tunu) *burn*. In other cases, Gitua may have retained a final consonant that was lost in Kove, as with G. toŋor, K. toŋo *mangrove* (POC *toŋo but PAN *(tT)eŋe(rR)). In still others, however, Gitua has a final consonant that does not seem to derive from POC, as with G. manum *bird* (K. manu, POC *manu(k)) and G. novuŋ *stonefish* (K. nou, POC *n opu). The origin of these is obscure. Finally, in some examples involving word-final /k/ and /p/ in Gitua, there are problems of historical sequence of certain postulated shifts, which will be discussed below. These involve such correspondences as G. nanak, K. nanai *pus*.

These problems apart, the voiceless stops differ greatly in the regularity of correspondence between /t/, on the one hand, and /k/ and /p/, on the other. Gitua /t/ always = K. /t/, with a single exception: G. dui, K. tutui *straight*. As regards /p/, in word-initial position in Gitua it always corresponds to K. /p/. Examples are numerous, ranging from G., K. pa- 'causative prefix' to G. pudi, K. puri *banana*. Note, however, the case of G. van, K. pani *give* mentioned above. Medial /p/ in the two languages sometimes corresponds and sometimes not; shared cognates with /p/ in other than initial position are too few for a pattern to be discerned. Both languages contain tapu *fall* and lupu *gather together*, but G. tavi *wave to beckon* may be cognate with K. tapitapi *wave in the wind*. The data suggest that POC medial *p often became /v/ in Gitua and /∅/ or possibly /u/ in Kove. Examples include G. avu *lime* possibly cognate with K. yaua ~ eaua (POC *apuR); G. tavure, K. taule *conch shell* (POC *tapu-); G. livo, K. luo *tooth* (POC *lipon); G. livu, K. liu *sibling opposite sex* (POC *lipu); G. mavu *heal* (POC *mapo); K. saŋaulu *ten* (POC *saŋapulu); G. yavara *north-west monsoon wind*, K. awaha *rain* (POC *paRa(t));⁵ G. raravia, K. lailai *afternoon* (POC *Rapi), etc. The suggestion that *p sometimes became K. /u/ derives from the fact that this certainly happened to initial *p, but only when followed by *a, as in K. wanawana *hot* (POC *pana(s)) and K. wala *shoulder* (POC *paRa). If the same shift occurred medially before other vowels, it might explain the unexpected /u/ in the Kove word for *tooth*; it might be that a /u/ derived from *p assimilated a preceding *i. There are exceptions to the pattern just noted, in which *p is reflected as /p/ in each language: G. ipi *Tahitian chestnut* (POC *qipi(l)); K. manipinipi *thin, flimsy* (POC *manipi(s)). It is impossible to say whether G. yap *fire* (beside K. eai, POC *api) retained an original stop because it had become word-final by the dropping of a vowel before medial *p shifted to a fricative, or whether, since Gitua words do not end with fricatives, a fricative became a stop (again) when the vowel was dropped and its position became word final (but see below).

Similar questions arise concerning Gitua /k/. It often appears initially, but very few of the words have Kove cognates. The few include G. kokoleŋ, K. koko *mushroom* (POC *koko)⁶ and G. kwaro, K. kaho *scratch* (POC *kadu). This latter will be discussed below, with other labialised Gitua stops. In one or two cases cognacy is uncertain, as with G. kokopwariŋa *spherical*, K. kapo *round*. In several cases, G. /k/ corresponds to K. /∅/: G. kosi *widower*, K. osiosi *widow*

(see below); G. kukudu, K. uru *carry on the head*.⁷ In at least one case G. /k/ = K. /ɣ/: G. kikil, K. ɣiliyili *tickle* (POC *kidi). Although both languages contain words with medial /k/, none of those in Gitua has a cognate in Kove. A particularly interesting correspondence is that between a set of Kove words that end in -ai while the Gitua cognates end in -ak: G. lawak *spider*, K. lauai *spider-web* (POC *lawa); G. nanak, K. nanai *pus*; G. watak, K. watai *know*; and possibly G. rak, K. hai *south-east wind*. The final /k/ in the *wind* term raises the suspicion that we have here a Gitua innovation, since no other languages in the region show a medial or final consonant in this word. Tuam-Mutu, however, has watagi *know*; the POC word for *pus* has been reconstructed as *nana(q); and the PAN word for *spider* contains a final glottal. It seems possible, then, that this Gitua final /k/ represents a final or medial proto-consonant that has dropped out in Kove. There are, however, other examples of final /k/ in Gitua that correspond to a consonant in Kove: compare G. wuzak, K. uraye *knife*. With the *knife* word and a few others, Lincoln notes that the "newer form" substitutes /l/ for /k/, a shift he does not explain.

I had mentioned in 1973 that POC *k had a variety of reflexes in Kove, though /Ø/ was the most common (Chowning 1973:198-199). POC *k is lost in many common words: ani *eat* (POC *kani); lalao *go* (POC *lako); rae *go up* (POC *nsake); oto *extend* (POC *koto). It is, however, retained as /k/ in an equally large number of equally common words: kaho and karisi *scratch* (POC *ka(dr)u and *kari(s)); kuku *mussel* (POC *kuku); kukururu *thunder* (POC *kududu); kulikuli *skin* (POC *kulit); koso *shell a coconut* (POC *koso); etc. In two words from basic vocabulary, *k is reflected by /h/: iha: *fish* (POC *ikan) and hihiu *tail* (metathesis of POC *iku). A doublet of ani *eat* appears in two common compounds: haniŋa *food* and pahani *feed*. The differences between Kove and Gitua appear not in the absence of Kove reflexes of *k, but in the fact that in some Kove words it is reflected by /Ø/ where Gitua has a consonant in the cognate word. Examples, in addition to G. gani *eat*, include such basic words as G. lago *go* beside K. lalao and G. bega *defecate* beside K. vevea (POC *pekas). Many other examples of Gitua /g/ where Kove has /Ø/ appear in words not reconstructed for POC, such as G. mago, K. mao *no, not*. It should also be noted that Gitua sometimes has an initial /g/ where no consonant has been reconstructed for POC, as with G. geno *lie down* (K. eno, POC *eno) and G. gunu *drink* (K. unu, POC *unu).

As I pointed out earlier (1973:199), POC *q is usually reflected as /Ø/ in Kove; the exceptions are so few and so varied that they may all represent borrowings. As with *k, Gitua often shows a /Ø/ reflex in many cases, such as ate *liver* (K. ateate, POC *qate), but /g/ in a substantial number of others, such as gumwa *work in gardens* (K. umo, POC *quma) and pugu *base* (K. pu, POC *puqu). There are other reflexes as well, notably G. witi *penis* (K. uti, POC *quti), a peculiarity to be set beside G. wili *skin*. Occasionally *q is reflected as G. /k/: G. kosi *widower* (K. osiosi *widow*; POC *qosi *widow*; POC *qosi *mourn*).

Before leaving the voiceless stops, it is necessary to discuss the labialised stops of Gitua. Lincoln suggests (1976d) that /kw/ represents a borrowing from NAN languages. Kove lacks labialised consonants, but in a number of cases POC *a following a labialised consonant in POC has become K. /o/, as in K. mota *snake* (G. mwata, POC *ŋmata). (For other examples, see Chowning 1973:199.) Several cognates which are not at present attributable to POC reflect this correspondence between Gitua and Kove: G. bwae, K. voe *armpit*; G. bwale, K. vole *hip*; G. bwaro, K. voho *egret*; G. mwai, K. moi *taro*; G. mwalik, K. moli *curl up*; G. apwa, K. apo *belly*. Exceptions include the following: G. kwaro, K. kaho *scratch* and G. mwatotol, K. matolutolu *thick*, where the Gitua form is unexpected in the light of the POC evidence; G. rumwa, K. luma *house*, where the POC evidence

supports the Gitua form; and G. damo, K. ramoha *forehead*, where we might expect G. *damwa.

The nasals can be dealt with briefly. G. /m/ always = K. /m/, and in the great majority of words, G. /n/ = K. /n/. (For exceptions, see below.) Also G. /ŋ/ normally = K. /ŋ/. Lincoln's data show that G. /ŋ/ sometimes substitutes for /n/ both in borrowings (such as tiŋ *tin*) and in variant recordings of the same word. This may explain such irregular correspondences as G. ŋgoreŋa *grease*, K. gorena *cooked coconut cream*; G. gilaŋ *embers*, K. ŋilani *sparks*; and G. ŋuru, K. nuri *smell*.

If the two languages are descended from a recent common ancestor, we would expect the voiced stops of Gitua to represent the voiced fricatives of Kove (see below). The situation is greatly complicated, however, by the fact that Gitua also contains a set of voiced fricatives, just as Kove now contains a set of voiced stops. See Table 1:

Table 1: Phonemes of Gitua (from Lincoln 1977a) and Kove

	GITUA					KOVE		
voiceless stops	pw	p	t	k	kw	p	t	k
prenasalised stops	bw	b	d	ŋg*		b	d	g
nasals	mw	m	n	ŋ		m	n	ŋ
voiced fricatives		v(β)	z	g(ɣ)*		v(β)	r	ɣ
voiceless fricatives			s				s	h
liquids			l	r				
glides		w		y		w		y?
vowels		a, e, i, o, u				a, e, i, o, u		

Both Lincoln and I have difficulty detecting the difference between g (his ŋg) and ɣ (his g), especially in word-initial position; this may account for some apparently irregular correspondences.

* Gitua /ŋg/ (Lincoln's) and Kove /g/ are pronounced alike, as are Gitua /g(ɣ)/ and Kove /ɣ/, which should be clearly realised as one reads on.

The voiced stop (Lincoln's ŋg) in Gitua is typically reflected by Kove /ɣ/: G. ŋgal, K. ɣali *poke, pierce*; G. ŋgaru, K. ɣahu *bite*; G. ŋgin, K. ɣini *wait for*; G. ŋgaya, K. ɣaia *pig*; etc. But in addition, G. /g/ often = K. /ɣ/: G. gilaŋ *embers*, K. ɣilani *sparks*; G. giramu, K. ɣilamo *slit gong*; G. guvi, K. ɣuvi *heel*; G. bage, K. vayeveye *wing*; G. dogi, K. royī *betel pepper*; G. zige, K. riɣeriye *edge, border*; etc.

In fact, Gitua /g/ has three different correspondences in Kove: /h/ (the rarest, but found only in words where it reflects POC *k or *q); /ɣ/; and most often, as was indicated above, /Ø/. If the examples with /h/ did not exist, we could assume that POC *-k- and *-q- fell together in some language ancestral to both Kove and Gitua, and that this single proto-phoneme was retained as G. /-g-/ while disappearing in Kove-Bariaī. All the words with /h/ are in basic vocabulary,

but so are many that lack it. A possible explanation is mixing of two closely related languages, one of which had lost reflexes of the medial consonant and one which retained a fricative reflex of it, possibly unvoiced. The cases in which Gitua initial /g/ reflects POC * \emptyset may reflect borrowing from another direction. Where POC *k is reflected by K. /k/, the Gitua cognate sometimes has a different consonant; see G. nganga *finger*, K. kaka *little finger* and G. gururuŋ, K. kukururu *thunder*.

The overall pattern of irregular correspondences involving the stops is repeated when other phonemes are examined. I shall not present all the data here, but simply point out some examples. Usually, G. /l/ = K. /l/: G., K. la *go*; G. labe, K. lavelave *testicles*; G., K. lio hang *oneself*; G. loŋon, K. loŋoni *hear*; G., K. lua vomit; etc. But G. /l/ also = K. /h/: G. laya, K. haia *ginger*; G. tola *a cold*, K. toha *cough*. In one very common word, G. /l/ = K. /n/: G. lam, K. nama *come*. (On the other hand, sometimes G. /n/ = K. /l/: G. manino, K. malilo *calm weather*; G. nima, K. lima *hand*; G. ne/ni, K. le 'neuter passive marker'.)

G. /r/ = K. /l/ in many words but K. /h/ in an equally large number, and the difference is not clearly ascribable to derivation from different proto-phonemes (but see below). Examples of the first correspondence include G. pera, K. pela *open out*; G. par, K. pali *stingray*; G. puro, K. pulo *red paint*; G. rau, K. launi *leaf*; *hair*; G. ririu, K. liliu *bathe*; G. rumwa, K. luma *house*; G. sarum, K. salumu *needle*; etc. The second correspondence is found in G. nora, K. noha *yesterday*; G. rua, K. hua *two*; G. suru, K. suhu *liquid*; G. wariza, K. wahira *day before yesterday*; G. waro, K. waho *vine*, etc.

I hope that this material at least makes clear why I became worried about Lincoln's and my own earlier assumption that Kove and Gitua belonged to the same subgroup.

Although most of the correspondences between the five vowels in each language are regular, a number of exceptions can be found. For example, G. /o/ = K. /a/ in G. loŋo, K. loŋa *enter* and probably in G. polelek *revert to type* (see also G. pulelek *turn*), K. palele *turn*; *change into*, but the opposite correspondence is shown in G. alimaŋa, K. alimaŋo *mangrove crab*. Although Gitua agrees with Kove in having /o/ for the more usual /u/ in the word for *older sibling same sex* (G. toga, K. toa), the vowels differ in the word for *spouse* (G. azua, K. aroa), and also in the word for *slit-gong* (G. giramu, K. yilamo). But compare G. amora, K. amuhua *you two*. Still another irregular correspondence appears in G. musilan *softly*, K. misilani *slowly, carefully*, and also in G. tibu, K. tuvu *kin two generations removed*. In the latter case, however, Gitua agrees with Bariai, though not with Kaliai. As regards the correspondence G. anutu *God* with K. anitu *spirit*, Lincoln suggests (personal communication) that anutu may be a loan from Yabêm.

1.4 Isoglosses

Lincoln's Rai Coast wordlists strongly confirm the lexical unity and distinctiveness of what he later (1976d) called the Nĕro subgroup (written Ngero in Wurm and Hattori 1981) composed of three languages: that of Malalamai, Boŋa, and Yara; Gitua; and that of Tuam-Mutu, Malai Island, and the "small Siassi Islands". They will hereafter be called Malalamai, Gitua, and Tuam-Mutu (T-M). These languages share many lexical items which either do not occur among their neighbours or which, more rarely, are phonologically distinctive (see Ross 1977:table 10).

In the following discussion, a word will be called Ngero if it occurs in at least two of these three languages and in no others of the Rai Coast region extending from Sio to Ham. Unless there is something distinctive about their shape, I have excluded words derived from POC. The majority of the distinctive Ngero words have cognates in Kove or, if not, in Kaliai or Bariai. Because my data on Kaliai and Bariai are very limited, I may well not know about cognates that exist there. I should note that while I was correct in pointing out forms shared by Kove with Gitua and Malalamai in Chowning 1973:209, some of these are more widespread on the Rai Coast, and so do not set off a small subgroup (see below). The following list follows Lincoln's sequence and his spelling. The bilabial fricative is indicated by /v/. A dash indicates a non-cognate.

Table 2

ENGLISH	MALALAMAI	GITUA	TUAM-MUTU	KOVE	COMMENTS
<i>one</i>	-	eze	es	ere	
<i>four</i>	paŋe	paŋe	paŋ	paŋe	
<i>sole of foot</i>	-	age-lolo	age-lolo	ahe ai-lolo	
<i>breast</i>	tuzu	tuz	tus	туру	
<i>back</i>	dume	dume	dimo	rumerume	K. back of crab only
<i>head</i>	dawa	dava	daba	Kal.rava	K. ravarava <i>top of tree</i>
<i>blood</i>	siŋ	siŋ	siŋ	siŋi	
<i>arse</i>	bole	bwale	bole	volevole	K. <i>hip</i>
<i>drink</i>	unu	gun	un	unu	Initial vowel distinctive in this region
<i>smell</i>	nuŋuzi	ŋuzu	-	nuri	
<i>hear</i>	loŋon	loŋon	loŋ	loŋoni	Final consonant unexpected
<i>know</i>	wote	watak	watagi	watai,atai	
<i>laugh</i>	ŋiŋ	ŋiŋ	ŋiŋ	ŋiŋi	
<i>sister's son</i>	-	waga	waga	waha	
<i>sibling opposite sex</i>	livu	livu	livu	liu	
<i>spouse</i>	azuwa	azuwa	azo	aroa	
<i>name</i>	eza	ezaŋgani	iza	era	
<i>spider/web</i>	-	lawak	lawak	lawai	
<i>wing</i>	bae	bage	bage	vaye	
<i>high tide</i>	-	sir	sir	sili	
<i>betel pepper</i>	doi	dogi	dog	royi	
<i>Saccharum edule</i>	tambol	tambuar	tambogar	tavuahi	Retains final consonant
<i>alight, of fire</i>	yan	gan	gan	ani	Same word as <i>eat</i>
<i>red</i>	-	sisinŋia	siŋsinŋia	siŋisinŋia	See <i>blood</i>
<i>digging stick</i>	wazo	yazo	yaz	waro	Bariai earo
<i>tie</i>	-	bitu	biti	piti	
<i>plait</i>	wowe	wowai	wawai	wawai	
<i>needle</i>	-	sarum	sarum	salumu	Retains final consonant
<i>split</i>	sosa	sasa	rasa	sasahai	Final -i in K. is a passive marker

(Table 2 cont'd)

ENGLISH	MALALAMAI	GITUA	TUAM-MUTU	KOVE	COMMENTS
<i>not</i>	mau	mago,ma	mako	mao	
<i>water</i>	ieu	yau	ya	eau	
<i>knife</i>	busai	wuzak	buza	uraye	
<i>hit</i>	lop	rap	rabi	hau	
<i>when?</i>	ŋeza	ŋeza	ŋez	ŋera	
<i>come</i>	nam	lam	la	nama	
<i>dance</i>	tol	tor	-	tohi	
<i>fall</i>	tapu	tapu	tap	tapu	
<i>good</i>	poe	pwaya	poe	Kal.poia	

There are, of course, other forms that have no known cognates in New Britain. These include the word *Lincoln* used to designate the subgroup, Gitua ŋero *man* (Malalamai ŋelo, Tuam-Mutu ŋeo); words for *woman* in all three languages cognate with G. *livage*; and a word for *tongue* that seems to have undergone metathesis, as in G. *yama* (cf. K. *mae*). On the other hand, a few terms are found not only in the Ngero languages and Kove but also in Sel, spoken on Arop and in Sel, Seure, and Mur villages. These include the word for *stab*, all cognate with K. *ɣali*, and the word for *sing*, cognate with K. *vou*; these two are not attested in the other New Guinea languages. Very occasionally, however, a cognate with the Kove form appears only outside the Ngero languages, as with Biliau *wos smoke* (K. *vosu*), with no cognate forms in the rest of the list.

Because I was confining myself in the table to forms characteristic of the Ngero languages, it does not include examples that unite a form found in only one of them with Kove. These exist, of course, and include T-M mem *urinate* (K. *meme*), T-M kis *hold* (K. *kisi*), and T-M manipi *thin* (K. *manipinipi*). They do not, however, so frequently link these Siassi Island languages to Kove as to suggest that Tuam-Mutu is much more closely related to Kove than are Gitua and Malalamai. Links to the New Guinea languages include G. matauzi *afraid* (K. *matauri*), particularly interesting because of the unexpected final consonant, and G. muso *dirty* (K. *muso dirt*), with cognates in other mainland languages but not within the Ngero 'subgroup'.

1.5 Grammar

Despite these irregularities and others not mentioned here, the two languages unquestionably share many lexical items. Grammatically, they show more obvious differences. Many of these reflect what could reasonably be called the greater complexity of Gitua. I had earlier commented on the 'grammatical simplicity' of Kove (1973:218), and Thurston, speaking of Kaliai or Lusi (grammatically almost identical with Kove) says (1982:35) that it "has the typological simplicity associated with creoles and it is therefore reasonable to entertain the notion that there has been a pidginisation process in (its) recent prehistory". This theory will be considered later; for the moment, I wish only to point out that Kove lacks many particles that modify verb phrases in Gitua and has no substitute for them (see Chowning 1978 for more information on Kove). The most conspicuous of these is the Gitua future marker *na*; futurity in Kove can only be indicated by using a connective preceding the subject of the phrase and meaning roughly *then*. Kove also lacks the Gitua prefix *para* or *pa* which indicates reciprocal or plural action. Contrast Gitua *isirua ti pa-rap* and Kove *asihua ti*

hau-ŋa-ri *they two fight*.⁸ It is not certain whether Gitua rap and Kove hau *strike* are cognate,⁹ but it can be seen that the languages express reciprocal action differently. On the other hand they both have, as was noted, a causative prefix pa-. The structure of what Bradshaw (1979) calls the serial causative differs, however. Compare Gitua rap-mate-a *strike dead it* with Kove hau ŋa i mate *strike (it) and it dies*.

The pronouns are very much alike, with one set of suffixed and three sets of independent possessives. The set in Kove formed of to + object pronoun is, however, an alternative to the third 'neuter' set, whereas in Gitua the cognate form "indicates that the owner retains title" (Lincoln 1976d). The independent subject pronouns, as distinguished from predicate markers, differ a good deal, but they also differ between Kove and Kaliai. The most noteworthy differences in the Gitua and Kove pronoun systems is that Gitua lacks the third person singular subject marker and suffixed possessive whereas in this same person Kove lacks the direct object marker (except for reflexive action) and has the others. Gitua does not have the Kove peculiarity of prefixing only the third person singular in the inseparable possessives; all follow the noun. Kove lacks the Gitua practice of usually infixing the second person singular subject marker u after the initial consonant of the predicate verb. The negatives are alike and both follow the verb, but the words indicating completion of action are different. So are the conjunctions.

Prepositions show notable similarities in both form and function. Gitua has a postposed locative e that is used somewhat like, and may be cognate with, Kove iai, but there is nothing corresponding with Kove aia or aea, as Thurston writes it (see Chowning 1973, 1978). A Gitua preposition nengan *for* or *with*, similarly is used like Kove ŋa(ni), and again may be cognate, though I have only one other doubtful example in which G. /ŋg/ = K. /ŋ/ (G. nanger *horsefly* beside K. laŋolaŋo *fly*). The other prepositions, G. pan, K. pani, and G., K. toma are certainly cognate, although in Kove the latter is used only with things, not with people. As was stated at the beginning, the overall impression is just that Kove grammar is simpler than that of Gitua, with few fundamental differences.

2. THE SIASSI ISLANDS AND MIGRATION

The earliest description of the linguistic situation seems to have been made by Bamler, a Lutheran missionary who allowed Chinnery (1926) to publish some of his lexical data. These indicate that there were in the 1920s three distinct AN languages, two (Barim and Iangla) on Umboi Island (Rooke Island) and one on the small islands of Tuam, Malai, and Aramot. The wordlists clearly indicate that this third language is closely related to Gitua. Much more recently, an anthropologist working there reported the following. Kaimanga is spoken on east Umboi and on Sakar Island, and is said to have originated on Umboi near the headwaters of the Simban River. Betang-Paramot or Karanai is spoken on Barim (once called Paramot), Mantagen, Aronaimutu (where they are bilingual), Tolokiwa, and Arop, and also on the Rai Coast of New Guinea from Kaiwa to Roinji. It is "very similar" to the languages of Gitua and Sio. Patrol reports trace migrations from Barim to the mainland of New Guinea, but local myths have Siassi settled from the Rai Coast and from Tolokiwa Island. Mutu has two dialects, one spoken on Aronaimutu, Mandok, Mutumala, and Aromot, and one on Malai and Tuam. Mandok (where the anthropologist worked) is most closely related to Kilenge of New Britain, and most Mandok trace descent from Kilenge via Aromot, but they also have migrants from other parts of Siassi and from the Tami Island area, including

Malasiga on the mainland (Pomponio 1983:1-17). Pomponio's statements about linguistic divisions agree with those of Harding based on fieldwork almost 20 years earlier, except that Harding simply says that "the Siassi Islanders of Mandok, Aramot, Malai, and Tuam speak closely related dialects of the same language", and does not mention the bilingualism of Aronaimutu, which he groups with Barim. He also says that all three of the AN languages of this area "appear to be closely related to each other and to the mainland languages of Sio, Gitua, and the Rai Coast" (Harding 1967:122-123). No mention is made of connections with New Britain.

Hooley, in his first attempt to classify the AN languages of Morobe, proposes four languages for the Siassi region: Lukep (Tolokiwa and the northern tip of Umboi); Mangap (Sakar Island and eastern Umboi, agreeing with Pomponio's Kaimanga); Barim (Aronai and south-west Umboi); and Tuam (Mutu) ("Mandok, Malai, and Tuam Islands, southern Umboi, and small settlement on the north coast of Huon Peninsula") (Hooley 1971:100). On his map (p.96), the settlement is labelled "Mutu" and is separated by Sio and Malasanga from Gitua to the east. The principal difference between Hooley's list and the others is the separation of Tolokiwa from Barim. This need not concern us, since those languages are not so closely related to Gitua or Kove as are Tuam and Mutu. Hooley provides separate wordlists for these two dialects or languages, and Mutu seems the more closely related to Gitua. Since Hooley also provides lists for Gitua and for Maleu (a dialect of Kilenge) any reader can assess the lexical evidence, while keeping in mind that as with any such wordlists there are some errors. (For example, the Gitua list has the word for *mouth* in place of the word for *tooth*, and the phrase *hit me*, in which the final consonant of the verb is dropped, for *hit*.) The relation between Gitua and Mutu is evident not only in the sharing of many specific lexical items but in the shapes of words that are more widespread. Particularly noteworthy are the presence of initial consonants in words for *eat*, *sleep*, etc.; of medial consonants in *older sibling same sex*, *new*, etc.; and of final consonants in *fire*, *sugarcane*, *south-east wind*, and *pus*, where neighbouring languages usually show /Ø/. Other shared peculiarities include the initial /n/ in nima *hand* and the metathesis in yama *tongue*. On the other hand, Tuam and Mutu, though particularly Tuam, show a strong tendency to drop final vowels which are retained not only in Gitua but in other languages of the region (see below). A comparison of Mutu with Maleu does not support Pomponio's assertion that Mutu is most closely related to Kilenge. Indeed, had this relationship been obvious Hooley would not have been so uncertain about whether Maleu belonged with the Siasi languages (see Hooley 1971:92,104).

Harding's account of language movements in this region differs somewhat from Pomponio's. He states that "Gituans trace their origins to tiny Pore Island (near Mandok)", and so do the people of "Malalomai" (sic); both places are supposed to have been settled by the passengers on two "canoes carrying migrants" which got separated. Where the people were migrating to, or why, is not explained. Harding has also collected many accounts of canoes being blown off-course in the Vitiaz Straits area, so that canoes from Siassi end up at various places along the north coast of New Guinea and canoes from Sio end up in the Arawe Islands off the south coast of New Britain. He also repeats Parkinson's account of finding drift voyagers from the D'Entrecasteaux and from the Trobriands on New Britain (the latter in a Kilenge-speaking region) (Harding 1967:12-13).

In his Rai Coast survey, Lincoln (1976a) notes that Harding's story indicating that Malalamai and Gitua originated from Por Island "seems to be quite recent and is now often told in connection with the tower of Babel. But a Siassi - New Britain source for both Boja (=Malalamai) and Gitua settlements would make sense,

a more permanent version of the settlement reported by Hooley". Here he is referring to his failure to locate a mainland settlement of Tuam-Mutu speakers located west of Malasaŋa in Hooley 1971, and his speculation that the people were actually "Siassi traders waiting for the Rai [south-east] wind to abate so that they could return to Siassi", Lincoln goes on to note that "The distribution of the so-called Korap Subfamily - Sel, Siŋorakai, Malasaŋa, Arop, Lokep, and Barim - suggests that these too may be remnants of trading expeditions".

Before examining further the reasons for some of these movements, it is worth mentioning that the Kaliai also trace part of their origins to Siassi. Haddon reports that:

According to tradition there was trouble in the island of Tuam and half the population came across to New Britain ... made smaller paddling canoes, and eventually settled in Kaliai. Tuam and Kaliai have many words in common, and between them exists a very friendly attitude.

(Haddon 1937:154)

Dorothy Counts (1968:49-50) recorded a Kaliai myth in which the crew of a large Siassi canoe came to the Kaliai area and brought civilisation to the local people, some of the Siassis settling there. In addition, according to Thurston, "Michael Freedman was told a story in Siassi about a fight which resulted in two canoes departing with people who settled in Kaliai" (Thurston 1982:60). Freedman, like Pomponio, worked in Mandok. (The Kove, however, have no traditions of migration, but assume that they originated on the coast in the middle of their present region, which was otherwise unoccupied. Judging from the number of generations involved, the single village from which they all trace descent broke up and dispersed about 1800.)

We need not take any origin myth at face value, even when it does not involve fantastic elements. It does not seem likely that four languages (one NAN, three AN) actually originated on Umboi Island and spread from there far along the north coasts of New Guinea and New Britain as well as, in the case of the NAN one, deep into the Huon Peninsula (Harding 1967:13). On the other hand, there is no denying that this is a region of constantly shifting populations. Some of the movements have been accidental, as with the drift voyages mentioned earlier, while others involved purposeful migration. One reason has been natural disaster. In this region, a major factor has been volcanic activity, including associated tidal waves. The entire north coast of New Britain contains a chain of active volcanoes which extend west through the islands off New Guinea as far as the Schoutens. Eruptions have occurred frequently in historic times, and geologists trace many more to the very recent past. The worst known devastation in western New Britain was caused by the eruption of Ritter Island in 1888; the resulting tidal wave was 12m. high at the western end of the island (the present-day Kilenge-Maleu region), extending up to 1 km. inland and causing landslides as well as flooding. Parkinson says that: "Zahlreiche Dörfer der Eingeborenen wurden fortgeschwemmt, und ein grosser Teil der Bewohner muss ... das Leben verloren haben". A German-led expedition camped on the shore at the time was obliterated without trace (1907:30). Undoubtedly the devastation extended to the islands of Dampier Strait as well. The Kilenge-Maleu region is itself one of active vulcanism; the local cones "have a history of strongly explosive eruptions, the earliest of which was recorded towards the end of the last century" (Löffler 1977:78). There are no active volcanoes in the Kaliai-Kove region, but Löffler estimates that the last eruption on the "entirely volcanic" Willaumez Peninsula just to the east occurred "at the beginning of this century" (1977:78). From oral accounts we know that

when a volcano was erupting frequently near Cape Hoskins early in this century, most of the population had to leave because their gardens were wiped out by ash-falls. They moved both east and west, and many stayed away for years. Judging from Johnston's description (1980) of dialectical differences in "Nakanai", the speech of some of my informants had been considerably affected by these sojourns. More recently, eruptions on Arop (Long) Island have forced population shifts (Harding 1967:133).

As well as volcanic eruptions, droughts also caused famines, particularly in places dependent on taro, which needs abundant rainfall. Harding says (1967: 92): "During a particularly severe famine which is supposed to have occurred late in the last century, a number of (Sio) people were forced to migrate to the Rai Coast and to Siassi". Other reasons for dispersal were overcrowding of the smaller islands, internal quarrels, and attack from outside. All of these have affected the distribution of the Kove in New Britain, with internal dissension especially leading to the establishment of new settlements far distant from the old ones, and in one case completely outside Kove territory (see Note 3). Harding also tells (1967:179) of people migrating from Siassi to the New Guinea mainland because of a local quarrel.

The region which includes the Vitiaz Straits, extending along the north coast of New Guinea and both the north and south coasts of New Britain, is tied together by an elaborate trade network. In many places trade is carried out by individuals whom Harding calls "trade friends", and at least in Sio trade was not a reason to encourage intermarriage with actual or potential partners (Harding 1967:181). Nevertheless, such intermarriage is common, especially in the smaller groups; were outsiders not admitted, many people would have to remain unmarried. (In 1963 Gitua's population was 415, and Mandok's had expanded to 343 from 120 in 1911: Harding 1967:114.) Hooley mentions marriage and trading as possibly expanding the percentages of shared cognates between Tuam and Tami, as well as other pairs of Morobe languages (1971:100). Where people actually travelled from Siassi, they are said once to have used a "Pidgin form of the Siassi language" (presumably Tuam-Mutu) for trading purposes, but Harding also says that "multilingualism is characteristic of the area" (1967:6).¹⁰ Farther from the Vitiaz Straits, multilingualism seems to have been the only solution for those who wished to trade abroad, and it was sometimes accomplished by sending young men to live in foreign communities so that they could learn the language. A Kove-speaking settlement on Bali Island is said to have resulted from this practice; inevitably some of the men decided to marry and stay there.

What seems likely to have resulted from this frequent moving around by both individuals and groups is communities of much more mixed linguistic origin than is found in some other regions such as the interiors of some of the larger islands. Especially where the travellers settled in places that could not support a large population - and that is true of Gitua and Sio as well as of the small Siassi Islands (Harding 1967:114) - we should expect the linguistic impact of a small number of immigrants to be much greater than if a single canoe landed among a larger population.

3. NEW SUBGROUPING PROPOSALS

As was noted above, Lincoln has proposed removing Kilenge-Maleu from Bariai, while keeping Gitua. In doing this, he is rejecting the evidence of the high cognate percentages uniting Kilenge with Bariai proper (44%). In discussing

other Rai Coast languages (1977b), he states that it can be argued "that all of the high cognate percentages between Huon languages and Rai languages are inflated by borrowing", especially as regards specialist traders, notably the Tami. In general, the subgroupings proposed here tend in fact to ignore cognate percentages, which were also the grounds for proposing initially that Bali-Vitu belonged with the Willaumez languages and Tubetube with those of Normanby Island (see below).

3.1 Lincoln

Lincoln accepts that the six languages in his redefined Bariai are "very closely related", and "show ... relative cohesion as opposed to their quite closely related Oceanic neighbors". He suggests calling the division composed of Malalamai, Gitua, and Tuam-Mutu "eastern" and the one containing Bariai, Kaliai, and Kove "western", but does not give the grounds for the division apart from noting that the eastern languages have tuzu *breast* and siŋ *blood* while the western ones have turu and siŋi. (In fact, Bariai proper also has siŋ.) As will be seen, there is considerable phonological justification for this division, which I have accepted.

Because of its relevance to the general argument of this paper, it is worth examining some of the reasons for Lincoln's excluding Kilenge-Maleu from Bariai and assigning it to Sio. The reason is that the Kilenge dialects lack the "Bariai" words for *blood* (Kil. tepo); *no* (G. mago, K. mao, Kil. avo); *water* (G. yau, K. eau, Kil. iako); and *bone* (G. tua, K. tuatua, Kil. bolbo!). At the same time, it shares with the "Sio" languages which include the other AN languages of the Siassi Islands and Umboi, the use of a word which elsewhere means *wing* (G. bage) for *hand*, a reduced form of the word for *ear*, and a special word for *tooth* represented by Kil. roŋa. He notes that "Bariai and Kaliai appear to have borrowed" the *tooth* form from Kilenge, and that Bariai has also borrowed the *hand* form from Kilenge (Lincoln 1977b), but does not consider the possibility that Kilenge may also have borrowed these terms. The destruction after the tidal wave would have left much of the coast open to settlement from the islands as well as from the interior.¹¹ In addition, some of distinctive Kilenge words have cognates elsewhere in New Britain. Specifically, Amara, an AN language spoken just to the east of Kilenge has topo *blood* and also kono *sleep* (see Kil. kono beside K. eno). As Friederici pointed out, Bali-Vitu also have topo *blood*. The Whiteman languages farther east in New Britain have cognates of the Kilenge word for *no*; cf. Sengeng awo.

I have nevertheless been persuaded by Lincoln that Kilenge is by no means as closely related to Kove as is Gitua.

3.2 Ross

Ross has suggested on phonological grounds that Kove does not subgroup with Gitua, while admitting that the lexical evidence in Chowning 1973 "calls into question the interpretation" that would put Proto-Bariai into a separate branch of Proto-Siassi from Proto-North Coast, assumed to have given rise to Gitua and many other languages of the north coast of New Guinea. The specific differences he cites are the following: in Kove,

- (a) the medial reflex of POC *p is zero;
- (b) the reflexes of POC *mp and *ŋk are fricatives, where other languages have stops;
- (c) POC *k and *q are lost entirely;
- (d) the reflex of POC *ns/*nj is r;
- (e) POC final vowels are consistently retained.

(Ross 1977:55-56)

As I have pointed out above, (a) is probably not correct, apart from cases like manipinipi *thin* which I cite there; kapi- *take hold of* (POC *kapit); and tuva *Derris* (POC *tupa). The probability that *-p- was sometimes reflected as K. /u/ is most evident in K. *saua what?* (POC *sapa); see also uwe *taro stick* (POC *upe) and examples given above. Since some of the other North Coast languages have luwo *tooth*, which Ross assumes shows a /w/ reflex of *-p- (1977:16), we do not seem to have a significant distinction here. I have also pointed out that *k and *q are not always lost in Kove; examples of reflexes of the latter include K. *ahe leg* (POC *waqe) and *tahe faeces* (POC *taqe). Since the reflexes of *mp and *ŋk are stops in Bariai proper, and no one can possibly doubt that Bariai is the closest relative of Kove-Kaliai, point (b) seems to reflect a misunderstanding of what I said (1973:195) about the shifts that led to the development of fricatives only in the eastern branch of these languages. At least, the situation seems to be clear as regards reflexes of POC *ŋk in Bariai. Not only does Friederici consistently write g in words like waga *canoe*, but he contrasts the sound heard there with one he writes γ which he heard in only a few words (1912:171). (Ironically, the one of these that has a Kove cognate - Bar. *tayahau fish spear* - is one in which I have recorded both sounds.) The fact that Friederici also used only g in writing Kove words raises some question about the sound being represented, but it seems clear from his mentioning the prenasalisation of the Bariai word for *pig* (his *ngaia*; see Gitua *ŋgaya*) that it was not pronounced like the present-day Kove and Kaliai *yaia*.

The question of the reflexes of POC *mp is less clear. With the sound that he writes b, he speaks of hearing a v-like sound, so that he sometimes wrote the word for *tree* as *avei* rather than *abei*. His comparison of the Bariai with the Spanish "b=v" seems to make it clear that what he heard was a bilabial fricative. He also notes that "ein nachlässig ausgesprochenes Barriai -w ist nicht weit von einem -v entfert", but decided not to use that symbol because "ein v kennen die Barriai nicht" (1912:169). In recording Kove, he writes b or w where I would write v, but interestingly has Kove *awe tree* beside Bariai *abei*. The Nicholsons write Bariai *boŋ night*, but the Haywoods heard a bilabial fricative in both this word and the word for *betelnut*, both reflecting POC *mp. It is worth noting that according to the Haywoods, in Maleu (which borders on Bariai), voiced stops "occur only as allophones of voiced fricatives" (1980:48). If their recording of Bariai is accurate, perhaps a tendency which in Friederici's time was found only in the bilabial voiced stop has spread to other voiced stops under influence from the west, duplicating what happened in Kove and Kaliai to the east (see under 4.).

Although POC final vowels are consistently retained in Kove itself, this is not the case as regards *-i and *-u in Bariai (see Chowning 1973:196 and many examples in Friederici).

Nevertheless, Ross is right in suggesting that there are significant phonological differences between Lincoln's eastern and western Bariai, creating difficulties which I try to deal with below.

3.3 Wurm and Hattori (1981)

In the Madang Province map they published, with no indication of the source, the Bariai subfamily has been divided into the Cape Gloucester section, which contains Kilenge and Maleu, and the Kove-Bariai section, while the Ngero subfamily consists of Tami, Mutu, Gitua, and Malalamai (both subfamilies among many in the Siassi Family). On the basis of the wordlists in Hooley 1971, I cannot see the justification for assigning Tami to Ngero and, like Lincoln, I am not happy with the assignment of Kove and Gitua to separate subgroups, even though I am aware of the problems if they are put together.

3.4 Chowning

At the very least, I propose that for the time being, Kilenge-Maleu be considered less closely related to Bariai, Kaliai, and Kove than are Tuam-Mutu, Gitua, and Malalamai. In the subsequent discussion, I shall refer to these last three languages as Ngero - appropriately, since they alone contain cognates of that word for *man* - but shall exclude Tami. To avoid confusion with other uses of the term Bariai by myself and Lincoln, I shall call his "eastern Bariai" languages Kove-Bariai, with a recent common ancestor proto-Bariai.

4. PROTO-DAMPIER

If it is assumed that Gitua, Malalamai, Tuam-Mutu, and Kove-Bariai have a single common ancestor separate (at a low level) from that which gave rise to other languages in the vicinity, I shall temporarily label it Proto-Dampier, after the strait separating New Britain from the Siassi Islands. Given that many isoglosses connect and mark off these languages, the question is whether any plausible phonological history can be reconstructed to account for the sound correspondences.

As regards vowels, Kove is the most conservative, followed by Gitua. Both have retained straightforward reflexes of POC vowels which have sometimes been dropped in other Dampier languages, particularly Tuam-Mutu, or have shifted, particularly in Malalamai. Kove, however, also has new final vowels following POC final consonants. They do not reflect vowel harmony, but are usually /i/ except after POC *-m, where /u/ occurs. The reason is probably that word-final /u/ following /m/ in Kove is dropped in normal pronunciation. The few exceptions to the 'rules' just given include atunu *tuna* and pelaka *lightning* (POC *pilak); compare K. samani *outrigger*, launi *hair*, foliage, varuhi *pigeon*, salumu *needle*, tilomu *oyster*, etc. (see list in Chowning 1973:197). It is suggested that Proto-Dampier regularly had *-i after *p which sometimes, at least, was derived from POC *-p, and after *k, sometimes derived from POC *-k or *-q. At an early stage, then, Proto-Dampier would have had not just something like *yapi *fire* but also *(C)utupi *draw water*; not only *Reki *grass* (POC *Reqi - Ross 1977) but *nanaki *pus* (POC *nanaq). (The data suggest that the POC word for *spider/web*, *lawa, should be reconstructed with a final *q as in PAN, producing Proto-Dampier *lawaki.) The same shape is reconstructed for words not derived from POC, as with PD *wataki *know*. At least in certain sequences, what then happened was that the medial *k and *p dropped out in the eastern languages, leaving such forms as K. utui *draw water*, nanai *pus*, lawai *spiderweb*, lailai *evening*, and watai *know*. In Gitua and Malalamai, by contrast, the final vowel was dropped, leaving such

forms as G. nanak *pus*, G. watak *know*, and Mal. laplap *evening* (beside G. raravia). In Tuam-Mutu, the final *i is dropped in some cases but tends to be preserved in verbs, and the stop became voiced: T-M rabrab *evening*, watagi *know*. The dropping of the consonant in the ancestor of Kove-Bariaai is characteristic of many languages along the north coast of New Guinea, according to Ross 1977.

It is very difficult to deal with some of the other possible proto-phonemes because of the available data, which are especially full of inconsistencies as regards the recording of Bariaai. It seems to have been more conservative than Kove and Kaliai as regards some consonants, but the three wordlists available to me (Friederici 1912, R. and R. Nicholson 1966 (MS), and G. and I. Haywood 1980) differ among themselves particularly in the recording of /r/ and /d/. sometimes offering alternative spellings of the same word. The question of Friederici's recording of some consonant sounds has been mentioned under 3.2 above. As regards an /r/ phoneme, there seem to have been two separate ones which he wrote in the same way. The first and more common was often heard, and probably pronounced, as an alveolar voiced stop, and so usually written as d. His examples make it clear that this was an allophone of an r-sound (a flap?) in certain environments, notably before /-e/, but he gives several examples of recording the same word with both r and d (e.g. ada/ara *our*(incl.) (edible)) (1912:169-170). The sound heard here seems to have differed from both the prenasalised /d/ he recorded in the word for *fire* (diŋa, pron. ndiŋa) and another r-sound he had difficulty distinguishing from /l/. Tests convinced him that this last should also be written as r, but it is interesting that one set of words in which he heard it was the part of the dual pronouns which contain the word for *two*, although he did not hear it in the word for *two* itself. Kove has hua and Kaliai ŋua in both the numeral and the pronouns. (I have not recorded Kove cognates for other Bariaai words in which he says /r/ is like /l/.)

The Nicholsons in recording Bariaai distinguish a flap /r/ from a trill, though the Haywoods do not. Since the trill in almost every case, like the Kaliai one, corresponds to Kove /h/, this suggests that proto-Bariaai had two /r/ phonemes. The same may have been true of proto-Ngero; Pomponio says that her /r/ represents both a flap and a trill. The inconsistent recordings make it difficult to be sure that Bariaai has a /d/ separate from these, or did in the past; Bar. diŋa *fire* beside Kil. and Kal. riŋa indicate that it does now.

Apparently in PB, the reflexes of POC *R, *d, and *l all had two reflexes, a trilled /r/ and /l/. The reasons for the separation are wholly unclear; possibly it reflects language mixing at an earlier stage. The trill remained in Bariaai and Kaliai but became /h/ in Kove, falling together with another /h/. In at least a few words this represented POC *-k- and *-q-; in other words, also shared with Kaliai, there is no known POC form. The flap /r/ in PB seems to be derived from POC *ns, and possibly *nd and *nt. Apart from Pomponio's statement, it is interesting that Malalamai once had a voiceless /l/, now pronounced /r/, in many of the same words in which Kove has /h/, with /l/ in others. See, for example, Mal. Lua, K. hua *two*; Mal. Lowo, K. hoho *fly*, Mal. Lop, K. hau *hit*. There are exceptions, such as Mal. nola, K. noha *yesterday*, but it may be that proto-Ngero once had two distinct phonemes corresponding to the trilled and flap /r/ of PB, which fell together at least in Gitua.

It is necessary to reconstruct a simple *g for PD, which was sometimes derived from POC *ŋk. In the Ngero languages this produced a prenasalised voiced stop, at least medially, leading to G. /ŋg/, and in PB perhaps a stop that was only lightly or occasionally prenasalised. In Kove and Kaliai this shifted to a fricative, so that a dozen cognate pairs show K. /ɣ/ = G. /ŋg/. There are, however, so many words in which both Kove and Gitua have a fricative, including

the first person direct object pronouns *ɣau*, *ɣita*, *ɣai*, that despite its apparent rarity in Bariai, it is tempting to reconstruct a voiced velar fricative for PB and PD. Because Gitua also has /ɣ/ as the reflex of POC *-k- and *-q-, however, a set of irregular correspondences has resulted: G. /ɣ/ = K. /ɣ/ in one set of words, and usually K. /θ/ in another, while G. /ŋg/ also = K. /ɣ/.

This is not the place to try to reconstruct PD, especially because it is impossible to do a thorough job without better data on some of the languages. It may be thought that the phonological histories are simply too divergent for Ngero and Kove-Bariai to belong to the same subgroup, but if they do not, how are we to account for the lexical connections? Apart from Thurston's theory, to be discussed below, there seem to be several possibilities. The distribution of Kilenge-Maleu suggests that it has been in New Britain a long time, with many speakers living deep in the bush, whereas all the other languages being discussed here are confined to narrow coastal strips and offshore islands. Unless an ancestor of Kilenge or a now vanished language once spoken in the same region gave rise to the Dampier languages, then the ancestor presumably came from outside New Britain. The only possibilities seem to be the north coast of New Guinea or Umboi Island. I would suggest that New Guinea was indeed involved at some stage to account for the presence of the preposed genitive and the *-iaɪ* suffix in Kove-Bariai (assuming that they were not derived from the NAN language Kovai on Umboi). According to this scenario, the western end of New Britain was already occupied by the ancestors of the speakers of Kilenge, and there may have been then, as now, speakers of other AN languages on Umboi. The speakers of PD may have already been affected by these neighbours before some migrated to the small Siassi Islands (and later to Gitua and Malalamai) and others to the north coast of New Britain east of the Kilenge region occupied by the Kilenge. They presumably remained together for a while, during which PB lost the consonants retained by the Ngero languages, but gradually spread east, coming in contact with different groups already resident in these regions, but not losing their initial linguistic unity. Both trade and warfare, in the person of refugees, also kept them in contact with the Ngero speakers of the Siassi Islands, but this language was also altering because of mixing with migrants from other regions. Nevertheless, diversification proceeded until the colonial period, when it became somewhat reduced by increased travel and contact resulting from the abolition of warfare. The Kove say that most of their ceremonies were acquired from Bariai or farther west during this century, and also that dialectical differences within Kove are disappearing. Meanwhile the Gitua and Malalamai people were being affected by a different set of neighbours, and diversifying in different ways. I am assuming, as I did earlier, that these languages arrived in New Britain very late in its period of settlement. I do not think that most of New Britain was settled in one wave of AN speakers, as Grace suggests (1986), nor do I think that Kove-Bariai is different enough to be assigned to a completely separate branch of his North Coast languages, as Ross suggests (1977).¹²

A different scenario has been proposed by Thurston, and since it seems to have been accepted without question by some other linguists (e.g. Lynch 1981:109), it needs detailed discussion.

5. THURSTON

In a recent study (1982), Thurston has compared Kaliai (= Lusi) with the NAN language Anem, spoken in Kaliai territory (see Map 2), and discussed possible influences between these languages. In order for the basis of my criticisms to

be understood, several points need to be stressed. Kove and Kaliai are very closely related, particularly in basic vocabulary, so that both David Counts and I have called them dialects of a single language. Counts says (1969:3) that the two are mutually intelligible, but in fact they are different enough so that misunderstandings sometimes occur, according to the Kove. The reason seems to be that each has borrowed from different neighbouring languages. Apart from terms relating to ritual and sailing, which are said to come from Kilenge and Bariai, Kove has borrowed from Bakovi, just to the east, and probably from Bali-Vitu. (Many Kove claim to speak Bakovi and Kilenge.) By contrast, the Kaliai seem to have borrowed from the Lamogai languages and, according to Thurston, Anem. According to Dorothy Counts (1968:48,242ff), some of the Kaliai are descendants of speakers of Lamogai languages (Aria and Lamogai proper) who migrated to the coast, at least partly because of warfare.

Presumably because of the patterns of borrowing, Kove and Kaliai differ much more in phonology and in lexicon outside the basic vocabulary than in grammar. Grammatically they are almost identical, though it is impossible to be sure about some points which are not discussed by David Counts. Phonologically the greatest difference is that Kaliai contains many consonant clusters, both initial and medial, and many words ending in consonants. Kove contains no initial consonant clusters; almost no medial ones except when certain reduplicated forms are pronounced rapidly (for example, *natnatu children* becomes *natunatu* in slow speech); and so few words with final consonants that they almost certainly are borrowings (e.g. *asipel k.o. mask*). One reason for the difference is that Kove has not undergone a shift, manifested in Bariai as well as Kaliai, in which POC *i and *u are dropped in certain positions, producing such differences as K. anitu, Kal. antu (POC *qanitu) *evil spirit* and K. tina, Kal. tna (POC *tina) *mother*. In many cases cognates not obviously of POC origin are shorter in Kaliai and lack vowels that are present in Kove.

Thurston suggests that Kaliai phonology has been influenced by Anem, particularly as regards consonant clusters and final consonants (1982:56) - quite possible, but not applicable to Kove. He then discusses lexical borrowing, demonstrating persuasively that Anem has borrowed many Kaliai terms having to do with canoes and the sea. He suggests that Kaliai in turn borrowed heavily from Anem in two areas, vocabulary referring to plants and animals of the bush, and a series of verbs that begin with *kə-* in Kaliai and *ge-* or *gə-* in Anem, having to do particularly with sound or motion. To take the second group first, of almost 40 cognate verbal forms of this sort shared by Kaliai and Anem, only one appears in my Kove data, and it is not a clear-cut case: Kal. *kamuřu*, Anem *gemuxu coo*, *of pigeons*. This is one of only two pairs in which the Anem word ends with a vowel, raising a question about the direction of the borrowing. The possible cognate is Kove *kamuru whisper*, which has the same shape and meaning in Bariai; the Kove word for *coo* is *kukururu* (see POC *kudu *dove*).

As regards the bush vocabulary, as Thurston points out, there are problems of identification, especially because so many of his terms rarely appear with accurate designations on wordlists. My Kove vocabulary is also particularly deficient in this area, for reasons having partly to do with the great maritime orientation of the Kove. Kove does share many of the plant names on Thurston's list of Kaliai-Anem cognate pairs, but some of these are found outside the area - a possibility he acknowledges. The ones with characteristic Anem phonological shapes do not have Kove cognates, as far as I know. Kaliai certainly does share many words referring to the bush with Anem. Curiously, however, Thurston does not seem to consider the possibility that both languages may have borrowed from the Lamogai languages which are much more widespread in this region than are Anem

and Kaliai (see Map 2). My lists for Lamogai languages are very brief, and contain only one term that relates to these pairs of bush words: *hornbill*. This is Kaliai *meřiaŋ*, Anem *mexiaŋ*. In two Lamogai languages we find: Lamogai proper *meřiaŋ*, Mouk *mahiaŋ*. (The Kove word is not cognate.) That terms have been diffused is certain, but more languages need to be examined before the source can be identified. (Thurston does acknowledge this fact - 1982:80.)

The weakest part of the argument has to do with grammar. Thurston begins by reconstructing what he calls "Standard AN", drawing only on the syntax of languages spoken outside Melanesia, and then goes on to assume that features found in Melanesia that diverge from this pattern represent NAN influence. Most remarkably, he argues that SVO sentence order is the result of such influence (p.16). Other evidence has to do with "modalities marked with particles in clause-final position" in both languages (p.33). In fact, every one of the modalities he lists (which are not cognate in Anem and Kaliai) are found in the same position in Gitua, and so are the separable possessives (Lincoln 1976c). Furthermore, the virtual absence of preverbal particles and the infrequent use of prepositions in Kove, which led me to speak of its grammatical simplicity (Chowning 1973:218), seem to be shared with Malalamai, and other apparent peculiarities of Kove-Kaliai appear in other languages of the north coast of New Guinea. In short, the data simply do not seem to me to support Thurston's conclusion that Kaliai was formed by Anem-speakers imperfectly learning a Siasi language, pidginising it. He adds that "it is possible to speculate that Kove is the result of the same processes ... but with a substrate dialect of Anem different from that encountered by the Lusi" (Thurston 1982:61). I am not opposed to the idea of pidginisation, and agree with Thurston that some of the Dampier languages may have undergone a degree of it at an earlier stage (and elsewhere), nor am I opposed to the idea of NAN influence on AN languages. I think, however, that there is no good evidence to support the suggestion that Kove has a substrate of a language that differs so enormously (not least in grammatical complexity) from itself.

6. JOHNSTON

Johnston has recently called attention to another problem involving Kove. In attempting to subgroup the New Britain languages (Chowning 1969), I assigned Bali-Vitu, the languages of the French Islands, to the "Kimbe Family", though with some qualms. The reasons for doing so were not just cognate counts but isoglosses connecting these languages with those of the Willaumez Peninsula, and the shared grammatical feature of the postposed genitive. Later (1973) I argued that contrary to Milke's assumption, the Kimbe languages do not belong with Kove and its relatives in being linked to AN languages of New Guinea. Johnston has reanalysed the data (1981, 1982) with particular attention to Bali-Vitu, which he assumes to be the most phonologically conservative of the Kimbe languages. He has much new material collected by himself and Ross, and also points out a fact that I overlooked in arguing for resemblances between Kimbe and EO: the frequent retention of POC final consonants in Bali. Cognate counts ranging from 42% to 47% still support the link between Bali-Vitu and the Kimbe languages of the Willaumez Peninsula (Johnston 1982:62) (though I do not accept the cognacy of some items assumed to derive from Proto-Kimbe, such as Bali *voraka* and Lakalai *ua root*). At the same time, an examination of wordlists collected by Ross and grammatical data presented in Johnston 1981 makes me increasingly uneasy about the assignment of Bali-Vitu to Kimbe or any other New Britain subgroup. Many

isoglosses also link these languages with Kove, even though Johnston found Bali and Kaliai to be only 25% cognate. Furthermore, various features of the grammar are reminiscent of Kove and unlike the one Kimbe language I know well, Lakalai (Johnston's Nakanai). These include the use of the preposition *tamani* *with* and the position of the possessives. Without more knowledge of the languages of the Willaumez, I cannot exclude the possibility that Bali-Vitu indeed links the Kimbe languages with the AN languages of New Guinea, as Johnston suggests. It may be, however, that the apparent links only reflect heavy influence from both the Willaumez languages and Kove. The fact that Kove has also borrowed from a Willaumez language (Bakovi, otherwise Bola) further complicates the picture.¹³ We have historical evidence of migration back and forth between the Willaumez Peninsula and Bali-Vitu, resulting from volcanic eruptions, famines, and fights. I mentioned earlier the Kove settlement on Bali that resulted from trade. Before a final decision can be made regarding Johnston's argument, we need to sort out the effects of migration to and from Bali-Vitu. Meanwhile Ross has argued (1983) that Bali-Vitu forms an isolate that should indeed be separated from the other Kimbe languages but that neither Bali-Vitu nor the rechristened Willaumez group belongs with Kove and its relatives. If he is right, and I suspect that he is, cognate counts led both Johnston and me astray.

7. TUBETUBE

A final example from Milne Bay will indicate how more evidence can decide arguments. Tubetube was grouped by Lithgow (1976) with Normanby Island languages but by Ross, on phonological grounds (1981), with Suau. After examining word-lists, I wrote:

I find it impossible to separate Tubetube from the languages of south Normanby, on the one hand, and from Sariba, which certainly belongs with Suau, on the other. I do not have enough data to decide whether Tubetube constitutes a true link between Suau and the D'Entrecasteaux or ... an unclassifiable amalgamation.

(Chowning 1981)

Macintyre has helped settle the question by documenting the history of the settlement of Tubetube, which has a tiny population. This has recently included a 'colony' from Normanby and other migrants from Panaeati superimposed on an 'original' settlement from Suau via Ware. The picture has been complicated by the use of several mission languages, including Dobu, and "intermarriage (encouraged by missionaries) between people of convert communities". When obsolete words are taken into account, the principal ties do indeed seem to be with Suau (Macintyre 1983:40-44). How similar Tubetube looks to languages other than Suau seems to vary with the age, education and immediate ancestry of one's informants. The language of the Amphlett Islands continues to cause problems in classification for both Ross and me, and may reflect an equally complex history, the result of their location and dependence on trade.

8. CONCLUSIONS

Despite the strong lexical ties, it remains uncertain whether it is possible to derive both the Ngero and Kove-Baria languages from a single low-level common ancestor ("Proto-Dampier"). If they do not belong to a subgroup separate from

Ross's other "North Coast" languages, the large amount of shared lexicon, which so impressed Lincoln as well as myself, needs explaining, but if they do form a separate subgroup, there remain difficulties associated with their phonological divergence. It is worth emphasising that in some features, Kove-Bariai resembles other languages of mainland New Guinea even though it differs from Ngero. For example, many of the North Coast languages are like Kove-Bariai in losing reflexes of POC *-p- in words for *fire* and *sugarcane*. Ngero differs not only from Kove-Bariai and from most other North Coast languages but also from a much more widespread Melanesian pattern in having lost the reflex of POC *k- in the word for *skin*. In addition to a few of the lexical items mentioned above, there are other cases in which Kove-Bariai resembles some languages of mainland New Guinea, though not Gitua and Malalamai, rather than other languages of New Britain, the French Islands, or the Siassi Islands. The retention of the final POC consonant in the word for *foliage, hair* (Kove launi, with a doublet laulau *leaf*) is found in several New Guinea languages, including Jabêm. An interesting case is that of Kove-Kaliai (not Bariai) tamine *woman*, with its unexpected /m/. In discussing the putative subgroup which would include both Ngero and Kove-Bariai, Lincoln commented (1977b): "Because Sisano shares so few similarities with Bariai languages, we can probably consider the Sisano /tus/ *breast* to be a coincidental resemblance." It is interesting, however, that Ross (1977) reconstructed Proto-Siau *tamein(e) to account for 'women' forms in Sisano and its relatives. If there ever was a specific link between Kove-Bariai and the AN languages of the West Sepik, it is unlikely to have been a strong or recent one, but these shared forms may be other clues to population movements in the region.

I have not considered here the considerable differences within the Ngero languages, which point to strong influence from other Siassi Island languages, in the case of Tuam-Mutu, and other mainland New Guinea languages, particularly in the case of Malalamai. The effects of immigration and intermarriage throughout the region, usually coupled with the effects of settling near other people speaking quite different languages, have obscured and confused what may once have been close relations among the languages spoken by some of the ancestors of these wanderers.

If my interpretation is correct, the relationship between languages can become much more complex than the usual tree model indicates, particularly because in some regions the paths of migration, deliberate and accidental, are so tangled. All of the historic evidence suggests that a considerable stretch of the north coast of New Guinea, including the offshore islands, has been subject to constant movements of people. Because so many different languages are involved and because some of them are closely related, the results are much more difficult to disentangle than, for example, in some parts of Polynesia. The relation between Gitua and its neighbours, including the languages of the Siassi Islands, is probably better represented as a series of overlapping circles than as a tree. Which part of the circles constitutes the core may be very difficult to ascertain. I feel dubious about the validity of several of the larger subgroups that have been proposed for both the Madang and Morobe regions (as in Hooley 1976, Z'graggen 1976) because the varying sorts of resemblances between languages assigned to the same subgroup suggest that the effects of population mixing and borrowing have not been distinguished from those resulting from descent from a single proto-language (see also Chowning 1973:209).

Of course, I myself have not succeeded in solving the same sort of problem as regards the relation between Kove and Gitua. A major impediment is inadequate information about the Siassi Island languages, particularly as regards phonology and grammar, but since there is every reason to think that their histories are

as complex as that of Gitua, uncertainties would probably remain no matter how much material was examined.

The example of Gitua and Kove, with their many irregular correspondences, indicates that more data about apparently related languages can simply raise new problems of interpretation. Where there has been so much movement and population mixing, perhaps it will never be possible to establish clear-cut subgroups. Probably many other parts of Melanesia have histories equally difficult to disentangle.

NOTES

1. Hooley and I have used the spelling with a single 's' for the language family as opposed to the islands, which are always Siassi. Others, particularly Ross, spell the languages like the islands.
2. The Kove material was collected in the course of my anthropological fieldwork, which was supported by the Australian National University (1966, 1968, 1969); the University of Papua New Guinea (1971-72, 1972-73, 1975-76); and Victoria University of Wellington (1983). Some linguistic work was also done when I visited Kove in 1978 on behalf of the Papua New Guinea Department of Environment and Conservation.
3. This is the western dialect spoken in the main Kove region, including the islands of Kapo and Nutanuvua, and not that of the breakaway villages of Tamoniai and Arumigi located far to the west. The Haywoods make a distinction between "Kove" and "Kombe" which wholly disagrees with my experience of people talking about the language rather than geographical divisions; the language (except when Pidgin is being spoken) is always called Kove, but the name also designates a few villages near a spot of that name, in about the centre of the region. Until recently all Kove called themselves Kombe in dealing with outsiders, but their recently established local government councils are called Kove East and West. The Haywoods' wordlist has under "Kombe" essentially the dialect I am describing, reserving "Kove" for the dialect of the breakaway villages which according to them exhibits some surprising phonological shifts: they write "Kove" th and ch as corresponding to "Kombe" r and ʎ (Haywood and Haywood 1980:46, 61-67).
4. The consequence is that I write either /w/ or /u/ to represent what in some cases is certainly a single phoneme deriving from POC *p. See Chowning 1973:238.
5. The comparative evidence indicates that the POC form should have a preceding syllable *qa-.
6. This term was reconstructed by R. French-Wright.
7. That a term with this meaning should be reconstructed for POC is indicated by Lakalai hugu, Motu udu.
8. In fact, both Bariai and Kaliai have a verb *to fight* which is pañao (Counts 1969:154) but Thurston specifically says that Kaliai lacks the reciprocal prefix and forms reciprocals like Kove (1982:27).

9. It seems highly probable that they are cognate, but there seems to be some uncertainty about the derivation. Bradshaw (1979) suggests POC *(dR)apat.
10. Haddon, however, thought that the principal trade language in north-west New Britain was Kilenge (1937:154).
11. Haddon says the coastal Kilenge villages were devastated by a smallpox epidemic - presumably that of 1896-97 - after which many hill people moved to the now thinly populated coast (1937:154).
12. Another possibility, implied by the various stories of migrations, is that the resemblances between the Kove-Bariai languages and Gitua and Malalamai simply reflect migrations from the small Siassi Islands to both New Britain and New Guinea. This would be more persuasive if first, Tuam-Mutu did not so clearly subgroup with the New Guinea languages rather than the New Britain ones and second, if the Bariai and Kove traced any of their origins to Siassi; they do not.
13. Bakovi and Bali-Vitu have so many words in common that although it is frequently possible to identify borrowings by Kove, on either phonological or distributional grounds, it is less often possible to say which of those languages was the donor.

REFERENCES

BIGGS, Bruce

- 1965 Direct and indirect inheritance in Rotuman. In G.B. Milner and E.J.A. Henderson, eds *Indo-Pacific studies* I:383-415. Amsterdam: North-Holland.

BODROGI, Tibor

- 1961 *Art in north-east New Guinea*. Budapest: Hungarian Academy of Sciences.

BRADSHAW, Joel

- 1979 Serial causative constructions and word order change in Papua New Guinea. *Papers in Linguistics, University of Hawaii* 11/2:13-34.

CHINNERY, E.W.P.

- 1926 *Certain natives in south New Britain and Dampier Straits. Territory of New Guinea Anthropological Report 3*. Melbourne: Government Printer.

CHOWNING, Ann

- 1969 The Austronesian languages of New Britain. *PL*, A-21:17-45.
- 1973 Milke's "New Guinea cluster": the evidence from northwest New Britain. *Oceanic Linguistics* 12:189-243.
- 1976 Austronesian languages: New Britain. In Wurm, ed. 1976:365-386.

- 1978 Comparative grammars of five New Britain languages. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, 1129-1157. PL, C-61.
- 1981 The Austronesian languages of the Massim. Paper presented at the Second Kula Conference, University of Virginia. Mimeo.
- COUNTS, David R.
- 1969 *A grammar of Kaliai-Kove*. *Oceanic Linguistics* Special Publication 6. University of Hawaii.
- COUNTS, Dorothy A.
- 1968 Political transition in Kandoka village, West New Britain. Ph.D. dissertation, Southern Illinois University.
- DARK, Philip J.C.
- 1974 *Kilenge art and life*. London: Academy Editions.
- FRENCH-WRIGHT, R.
- 1983 Proto-Oceanic horticultural practices. MA thesis, University of Auckland.
- FRIEDERICI, Georg
- 1912 *Beiträge zur Völker- und Sprachenkunde von Deutsch-Neuguinea. Mitteilungen aus den Deutschen Schutzgebieten*, Ergänzungsheft 5.
- GRACE, George W.
- 1986 Further thoughts on Oceanic subgrouping. In this volume, 1-12.
- HADDON, A.C.
- 1937 *The canoes of Melanesia, Queensland, and New Guinea*. (Vol. II of A.C. Haddon and James Hornell, *Canoes of Oceania*). Bernice P. Bishop Museum Special Publication 28. Honolulu.
- HARDING, Thomas G.
- 1967 *Voyagers of the Vitiaz Strait: a study of a New Guinea trade system*. Seattle: University of Washington Press.
- HAYWOOD, Graham and Irene HAYWOOD
- 1980 The languages and communities of north-western New Britain. In R.L. Johnston, ed. *Language, communication and development in New Britain*, 41-70. Ukarumpa: Summer Institute of Linguistics.
- HOOLEY, Bruce A.
- 1971 Austronesian languages of the Morobe District, Papua New Guinea. *Oceanic Linguistics* 10/2:79-151.
- 1976 Austronesian languages: Morobe Province. In Wurm, ed. 1976:335-348.
- JOHNSTON, Raymond L.
- 1981 Mimeographed version of Johnston 1982, with appendix on Bali.
- 1982 Proto-Kimbe and the New Guinea Oceanic hypothesis. In Amran Halim, Lois Carrington, and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, vol.1: *Currents in Oceanic*, 59-95. PL, C-74.

LINCOLN, Peter

- 1976a Rai Coast survey: first report. Mimeo. UPNG.
 1976b Rai Coast word list. Mimeo. UPNG.
 1976c Rai Coast Austronesian language survey data: Part II. Mimeo. UPNG.
 1976d Gitua basics, or Gitua varu pugu. Mimeo. UPNG. 10pp.
 1977a Gitua-English vocabulary: preliminary version. Mimeo. UPNG and University of Hawaii.
 1977b Subgrouping across a syntactic isogloss. Paper presented to the Austronesian symposium, Hawaii. Cyclostyled.
 1979 Proto-Oceanic cognate sets. Computer printout.

LITHGOW, David

- 1976 Austronesian languages: Milne Bay and adjacent islands (Milne Bay Province). In Wurm, ed. 1976:441-523.

LÖFFLER, Ernst

- 1977 *Geomorphology of Papua New Guinea*. Canberra: CSIRO and Australian National University Press.

LYNCH, John

- 1981 Melanesian diversity and Polynesian homogeneity: the other side of the coin. *Oceanic Linguistics* 20/2:95-131.

McELHANON, Kenneth A.

- 1970 A history of linguistic research in the Huon Peninsula, New Guinea. In S.A. Wurm and D.C. Laycock, eds *Pacific linguistic studies in honour of Arthur Capell*, 1178-1208. PL, C-13.

MACINTYRE, Martha

- 1983 Changing paths: an historical ethnography of the traders of Tubetube. Ph.D. thesis, Australian National University.

NICHOLSON, Ruth and Ray NICHOLSON

- 1966 Bariai word list. MS.

PARKINSON, Richard

- 1907 *Dreissig Jahre in der Südsee*. Stuttgart: Strecker and Schröder.

POMPONIO, Ali

- 1983 Namor's Odyssey: education and development on Mandok Island. Ph.D. dissertation, Bryn Mawr College.

ROSS, Malcolm

- 1977 Relationships of the Austronesian languages of the Sepik and Western Madang coast of New Guinea. Mimeo.
 1983 The genetic relationships of the Austronesian languages of Papua. Paper presented to the XV Pacific Science Congress, Dunedin.
 1986 A genetic grouping of Oceanic languages in Bougainville and the Western Solomons. In this volume, 175-200.

THURSTON, William R.

1982 *A comparative study in Anêm and Lusi.* PL, B-83.

WURM, S.A., ed.

1976 *New Guinea area languages and language study, vol.2: Austronesian languages.* PL, C-39.

WURM, S.A. and ShirÔ HATTORI, eds

1981 *Language atlas of the Pacific area, Part I.* Canberra: Australian Academy of the Humanities, in collaboration with the Japan Academy; PL, C-66.

Z'GRAGGEN, John A.

1976 *Austronesian languages: Madang Province.* In Wurm, ed. 1976:285-300.

WAI: A MALAITAN LANGUAGE IN FIJI

Jeff Siegel

0. INTRODUCTION

One of the more recent migrations of speakers of Austronesian languages took place between 1864 and 1911, when approximately 27,000 people from island Melanesia and Micronesia came to Fiji as indentured plantation labourers. Although most of those who survived returned to their home islands at the end of their contracts, a large number did stay on in Fiji. Some of their descendants have been absorbed into Fijian society, but the majority are still a distinct group who, without land rights, live together in settlements around the country.

This paper is about these immigrant contract labourers and the languages they brought to Fiji. It focuses on one particular language, called "Wai" by its speakers, which has survived to this day, although now it is on the verge of extinction. Wai has not been investigated thoroughly, but some preliminary information is presented here which may be of interest from the perspective of language contact and mixing. It may also provide some insights to those trying to reconstruct what happened in the more distant past.¹

The paper starts with some background information about the origins of the Pacific islands labourers in Fiji and about their descendants. The next section deals with the languages spoken by the labourers - especially those from Malaita, the homeland of the Wai speakers. Then in the main section, Wai is described in comparison to the North Malaitan dialects from which it appears to be derived. Finally, the findings are discussed in the context of the general sociolinguistic areas of language mixing and language attrition.

1. BACKGROUND

1.1 The origins of Pacific islands labourers in Fiji

The Pacific labour trade (described in detail in Parnaby 1964; Scarr 1967, 1973; and Corris 1973) involved the recruitment of men and women from island Melanesia and Micronesia to work in Hawaii, Queensland, New Caledonia, Samoa, and Fiji. Labourers for Fiji were obtained to work for European owners of cotton, copra, and sugarcane plantations in rural areas or of business concerns in the urban centres of Levuka (the first capital) and Suva. Later imported labourers also worked for the colonial government (see Kuva [1974]:10).

From 1865 until the end of the labour trade in 1911, an estimated 27,027 islanders became contract labourers in Fiji. They were imported from the areas

Paul Geraghty, Lois Carrington and S.A. Wurm, eds *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*, 435-463. *Pacific Linguistics*, C-94, 1986.

© Jeff Siegel

435

now known as Vanuatu, the Solomon Islands, the New Guinea Islands region of Papua New Guinea, and Kiribati. Figures by year and area are given in Table 1-1 (from Siegel 1982). The employment of Pacific islanders reached its peak in the early 1880s. Up to this time, the main recruiting area was Vanuatu, but from the mid 1880s until the end of the trade, it was the Solomons. In the last dozen years, from 1900 to 1911, over 90 per cent were from the Solomons, and more than 80 per cent were from the island of Malaita. Numbers from the different islands of the Solomons, arranged according to current political divisions, are given for 12 year periods in Table 1-2, along with percentages of the total (from Siegel 1985b).

1.2 Pacific islands settlers

Many of the labourers whose contracts had expired chose to stay in Fiji rather than return home. Those who did stay on were often men who were living with Fijian women. This was especially true of those from the Solomons and Vanuatu because the labour force from these island groups included only a small percentage of women - 3.2 per cent for the Solomons and 8.2 per cent for Vanuatu, compared to 41.1 per cent for Kiribati (Siegel 1985b:51-53). Also, most of the female labourers were already married before being recruited (Kuva [1974]:15).

Some of the islanders who stayed on lived among the Fijians and became "almost indistinguishable from them in appearance and speech".² But most of them either re-engaged as plantation labourers or obtained jobs in urban areas as house servants, road builders, storemen, or dock workers (Corris 1973:87).

Labourers who had finished their contracts began living together in settlements near urban areas. One of the first of these settlements was Koro Ivi near Nailaga, outside Ba (CSO 1029/1887).³ Like most of the settlements, it had a mixed population, consisting of men from Epi and Pentecost in Vanuatu and women from Kiribati.

Many settlements were located in the Suva area. One was at Nasinu consisting of men from Pentecost and the Solomons (CSO 5135/1905). (This settlement is most probably the still existing one called Manikoso, supposedly a corruption of the name Pentecost.) Two other early settlements in the Kaunikuila (Flagstaff) area of Suva were named after islands in Vanuatu: Malekula (Malakula) and Sadro (Santo) (Rabukawaqa 1967:6). Large numbers of Islanders also lived in Toorak, Tamavua, Vunidilo (Samabula), and Caubati.⁴

Other settlements were organised by the Anglican Church, which first started working with Melanesian labourers in 1870 in Levuka. At the church's suggestion that the landless ex-labourers live in a Fijian-style village, Naviro was built above Vagadaci in 1886 (Kuva [1974]:17). In the 1920s it was moved to Wailailai, Fijian-owned land leased by the church for the Islanders. In 1946, it consisted of 15 small houses and 150 people, and was described by one visitor as "a tropical slum - an Anglican slum".⁵

In 1940 the Anglican Church bought 254 acres of land at Wailoku, just outside Suva, to provide one large, centralised settlement for Anglican Melanesians from the Suva area and also from Ovalau, Vanualevu, Taveuni, and Rabe.⁶ It was officially opened in August 1942, and called the Patteson Settlement. Also in the 1940s, the Solomon Islanders who lived on Taveuni and in south-east Vanualevu in settlements at Maravu, Laulevu, and Vunilagi were encouraged to settle on the church property of Natoavatu Estate, about 24 kilometres west of Savusavu. In 1951, the settlement there called Naviavia was formally recognised as the Campbell Settlement.⁷

Table 1-1: Origins of Pacific island labourers in Fiji

year	Vanuatu	Solomons	New Guinea Islands	Kiribati	Other or unknown	TOTAL
1864	--	--	--	--	--	--
1865	180	--	--	--	--	180
1866	301	--	--	--	--	301
1867	264	--	--	--	--	264
1868	381	--	--	135	--	516
1869	262	--	--	136	--	398
1870	1,348	212	--	224	--	1,784
1871	1,569	336	--	54	316	2,275
1872	1,023	--	--	427	116	1,566
1873	1,232	--	--	34	--	1,266
1874	607	--	--	85	--	692
1875	185	--	--	66	153	404
1876	247	140	--	84	--	471
1877	332	155	--	50	--	537
1878	1,058	202	--	236	--	1,496
1879	1,335	468	--	70	--	1,873
1880	911	1,382	--	68	--	2,361
1881	763	464	--	--	--	1,227
1882	1,022	502	467	102	--	2,093
1883	273	339	662	276	--	1,550
1884	128	585	489	64	--	1,266
1885	50	245	--	--	--	295
1886	145	131	--	1	--	277
1887	70	175	--	28	--	273
1888	71	209	--	--	--	280
1889	62	46	--	--	--	108
1890	35	116	--	40	--	191
1891	79	153	--	136	--	368
1892	58	153	--	1	--	212
1893	--	--	--	--	--	--
1894	--	--	--	14	--	14
1895	24	115	--	67	--	206
1896	19	98	--	--	--	117
1897	--	--	--	--	--	--
1898	--	102	--	--	--	102
1899	8	92	--	--	--	100
1900	--	--	--	--	--	--
1901	7	62	--	--	--	69
1902	--	--	--	--	--	--
1903	17	105	--	--	--	122
1904	3	93	--	--	--	96
1905	12	103	--	--	--	115
1906	55	184	--	--	--	239
1907	62	502	--	--	--	564
1908	--	210	--	--	--	210
1909	--	361	--	--	--	361
1910	--	78	--	--	--	78
1911	--	110	--	--	--	110
TOTAL	14,198	8,228	1,618	2,398	585	27,027

Table 1-2: Origins of Solomon Islands labourers: islands and provinces

	1864- 1875	%	1876- 1887	%	1888- 1899	%	1900- 1911	%	TOTAL	%
Duff			1	--					1	--
Tikopia			1	--					1	--
TOT: TEMOTU			2	-					2	--
Makira			540	3.9	50	2.9	38	1.9	628	2.3
Ulawa			10	0.1	9	0.5			19	0.1
Uki			2	--					2	--
Santa Ana			6	--					6	--
S. Catalina			5	--					5	--
TOT: MAKIRA			563	4.1	59	3.5	38	1.9	660	2.4
TOT: GUADALCANAL			892	6.5	184	10.8	138	7.0	1,214	4.5
Malaita			2,694	19.6	826	48.6	1,593	81.1	5,113	18.9
Dai			1	--			3	0.2	4	--
Ontong Java			32	0.2					32	0.1
TOT: MALAITA			2,727	19.9	826	48.6	1,596	81.3	5,149	19.0
Savo			12	0.1			2	0.1	14	0.1
Gela			69	0.5	2	0.1	15	0.8	86	0.3
Bellona			4	--					4	--
TOT: CENTRAL			85	6.2	2	0.1	17	0.9	104	0.4
TOT: ISABEL			90	0.7	13	0.8	18	0.9	1,211	0.4
New Georgia			3	--					3	--
Ronongo			5						5	--
Vella Lavella			2						2	--
Mono			24	0.2					24	0.1
Choiseul			27						27	0.1
TOT: WESTERN			61	0.4					61	0.2
OTHER/UNKNOWN	548	5.7	368	2.7			1	0.1	917	3.3
TOTAL	548	5.7	4,788	34.9	1,084	64.8	1,808	92.1	8,228	30.4

Notes: Figures for 1870-71 most probably included the North Solomons. The 1 Other arrived in 1907, born in Queensland: father from Malekula, mother from Ambae. Percentages are of the total number of Pacific islands labourers indentured.

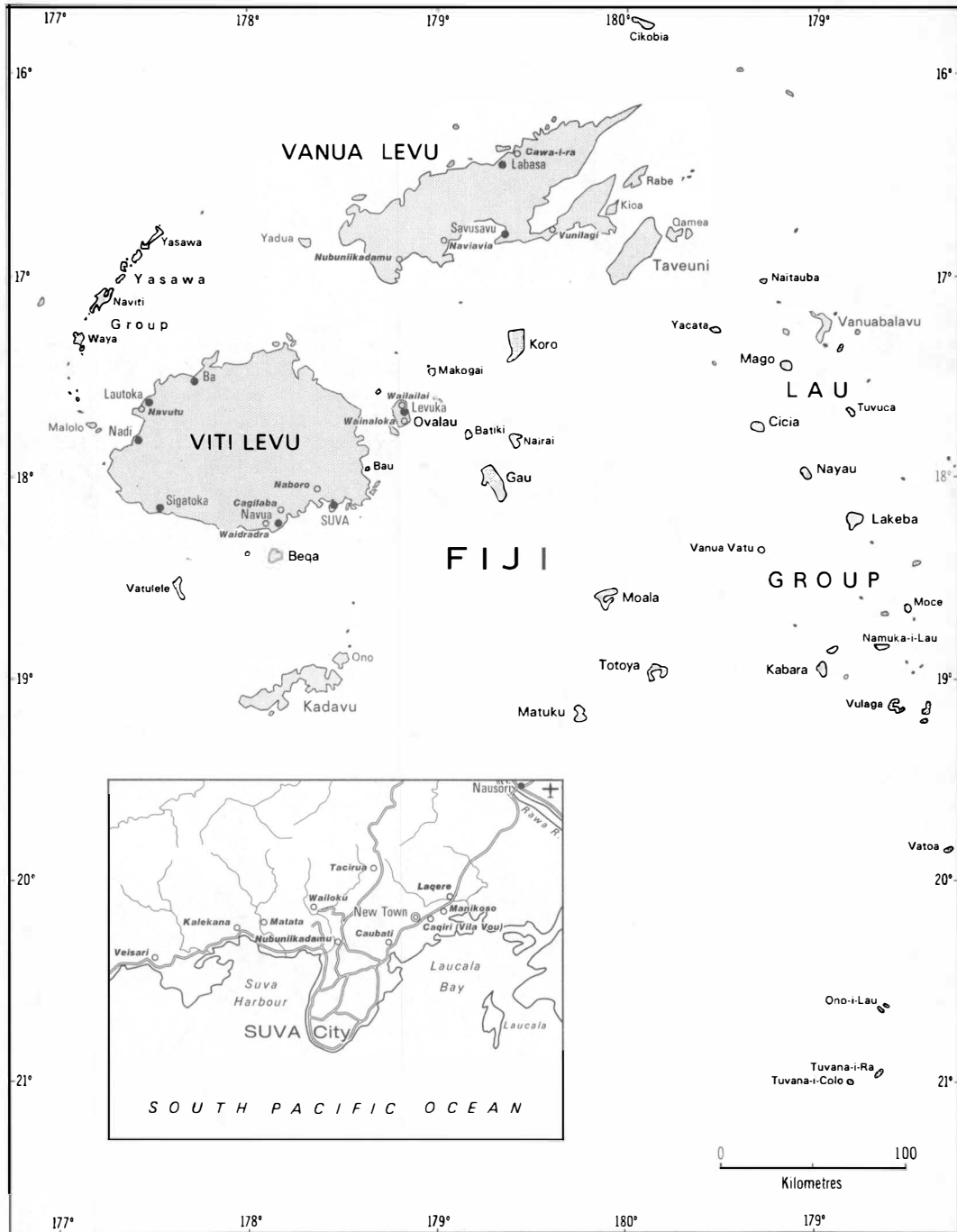
In 1952 a hurricane destroyed houses in many of the settlements around Suva, especially at Kaunikuila. The Anglican Church in cooperation with the Suva City Council established two new settlements, New Town at Nasinu and another at Lami (Kuva [1974]:18). The last new settlement was established in 1966 on a 465 acre block purchased by the church at Wainaloka, Ovalau, about 15 kilometres from Levuka.

1.3 Descendants of the labourers

According to the 1976 census, there were 6,822 "Other Pacific Islanders" in Fiji. In addition to the descendants of those who came as indentured labourers, these include free immigrants and their descendants from Tonga, Samoa, Kiribati, Tuvalu, and other islands. Nearly all the descendants of the labourers, however, are part Fijian, and many were probably included in the census with Fijians. Kuva ([1974]:19) gives the number of "Solomons" or *Kai Solomone* - as all descendants of Melanesian labourers are known in Fiji - as 8,000. He says that about 7,000 of these live in the four large church settlements mentioned above (Wailoku, New Town, Wainaloka, and Naviavia) and in other smaller settlements. Descendants of Kiribati labourers also live in some of the smaller settlements. A list of all the settlements is given in Table 1-3 with location, major island groups, principal religious affiliation, and approximate population (mainly from Rabukawaqa 1967). The locations are shown on Map 1. In addition, there are hundreds of descendants of Pacific islands labourers on Taveuni, formerly one of the major plantation areas.

During fieldwork in 1982 and 1983, I visited Taveuni and 15 of these settlements (marked with an asterisk in Table 1-3), and interviewed more than 30 descendants of the Pacific islands labourers. Most of these informants were older people who were of the first generation born in Fiji. Some of them are described in section 3.1.

I was also very fortunate to have been able to talk with the last surviving man of those actually recruited to come to Fiji during the labour trade. He was Jioji Abunio of Matata, Lami, originally from the Kwaikwaio area of Malaita. (He was also one of the original labourers interviewed by Kuva ([1974]).) Jioji died in December 1982.



Map 1: Pacific Islander settlements

Table 1-3: Pacific Islander settlements

NAME	LOCATION	MAIN GROUPS	CHURCH	POP.
VITILEVU: SUVA AREA				
* 1. Wailoku	NW of Suva	Sol	Anglican	600
* 2. Veisari	W of Lami	Kir	Protestant	200
* 3. Kalekana	Lami	Sol	?	?
* 4. Matata	Lami	Sol	Anglican	?
5. Tamavua-i-wai	Suva	Sol	Anglican	?
* 6. Tacirua	N of Suva	Sol	Methodist	?
* 7. Caubati	Nasinu	Sol	?	?
8. New Town	Nasinu	Sol, Van	Anglican	400
* 9. Caqiri (Vilavou)	Nasinu	Kir, Van	Catholic	100
10. Manikoso	Nasinu	Van	Anglican	140
11. Laqere	Nasinu	Sol, Van	Anglican	?
VITILEVU (OUTSIDE SUVA)				
*12. Navutu	Lautoka	Sol, Van	?	150
*13. Waidradra	Deuba	Sol	Anglican	150
*14. Cagilaba	Navua	Sol	Anglican	40
15. Naboro	Naboro	Kir	?	?
OVALAU				
*16. Wailailai	Levuka	Sol	Anglican	60
*17. Naisoqo	Levuka	Sol	Anglican	40
18. Korovou	Levuka	Sol, Van	?	?
19. Malekula	Levuka	Van	?	?
*20. Wainaloka	Lovoniiwai	Sol	Anglican	300
VANUALEVU				
21. Nubuniikadamu	Wainunu	Sol	Anglican	?
*22. Cawa-i-ra	Labasa	Sol	Methodist	60
*23. Naviavia	Wailevu West	Sol	Anglican	180
24. Vunilagi	Savusavu East	Sol	Anglican	?

2. LANGUAGES SPOKEN BY THE LABOURERS

Linguistic surveys of the Solomons and the other areas of the south-west Pacific (e.g. Tryon 1976; Wurm and Hattori 1981; Tryon and Hackman 1983) show that approximately 180 distinct languages are spoken in the islands from which the labourers originated. Only one group of islands, Kiribati, and a few of the smaller islands in other groups, have only a single language. The larger islands have several languages; for example, Santo and Malakula each have more than 25.

Since we know the islands of origin of the labourers, the number of speakers of each language who came to Fiji can easily be determined for the smaller islands. But for the larger islands it is impossible to estimate the number of labourers who spoke each language without knowing more precisely from which area of the island they originated. It is clear, however, that the language with the most speakers in Fiji was Gilbertese, spoken throughout the Kiribati group, which supplied 2,398 labourers. The next most important languages, especially in the later years of the labour trade, were almost certainly from Malaita, which provided 5,113 labourers, but from several distinct language areas. These are described in the following section.

2.1 The languages of Malaita

2.1.1 The current language situation on Malaita

Simons (1980, 1982) distinguishes between 12 "languages" spoken on Malaita based on "the language groups most generally recognized by Malaitans and in previous literature" (1980:3). These are given in Table 2-1 along with population estimates from the 1976 census:

Table 2-1: Malaita languages (based on Simons 1980, 1982)

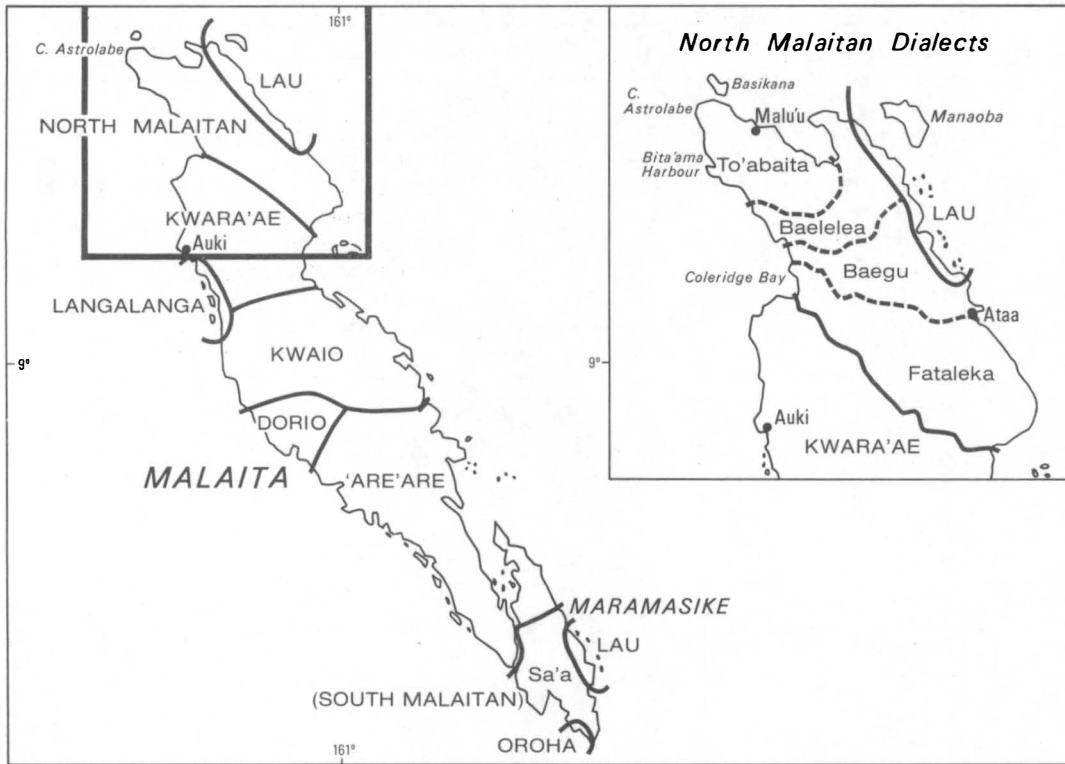
To'abaita	(TOB)	(5,226)
Baelelea	(BLE)	(4,252)
Baegu	(BGU)	(2,277)
Fataleka	(FTK)	(2,487)
Lau	(LAU)	(7,386)
Kwara'ae	(KWR)	(13,214)
Gula'alaa	(GUL)	(300)
Langalanga	(LNG)	(3,066)
Kwaio	(KWO)	(6,773)
Dorio	(DOR)	(571)
'Are'are	(ARE)	(7,225)
Sa'a	(SAA)	(4,445)

Simons points out, however, that if the criterion of mutual intelligibility were applied, there would be fewer languages, especially in the north where To'abaita, Baelelea, Baegu, Fataleka, and Lau are all mutually intelligible to some extent. This fact has been pointed out by several authors (e.g. Ivens 1930: 27), and in earlier studies, these five varieties have been grouped together as one language. It was called Lauic (Murdock 1964:120), and the encompassing language area referred to as Bali (Russell 1950:1; Ross 1973:49).

Simons (1980:3) also notes that if we recognise the distinct speech-culture groups that Malaitans themselves identify, there would be even more languages. The languages of small groups of this type have been called "communalects" in the Fiji context (Pawley and Sayaba 1971; Geraghty 1983) but the term could equally apply to Malaita.

In the most recent linguistic survey, Tryon and Hackman (1983) use the criterion of generally acknowledged mutual intelligibility to group together To'abaita, Baelelea, Baegu, and Fataleka as "major dialects or sublanguages" (p.27) of one language, North Malaitan. They also apply the criterion of sharing more than 80 per cent cognate basic vocabulary. The authors point out, however, that by applying the same criteria, Lau could have also been included in North Malaitan, but following tradition it was considered a separate language (p.27n). For convenience in this work, "North Malaitan dialects" will include Lau (which is also spoken in a small coastal area of South Malaita), and North Malaitan (NM) will refer to the linguistic system of which the dialects are subsystems.

Also in contrast to Simons, Tryon and Hackman group Sa'a along with the Ulawa and Uki Ni Masi dialects as South Malaitan. In addition, they use the name Kwai for Gula'alaa. Finally, they include another small language group, Oroha. The languages and major dialects of Malaita then, according to Tryon and Hackman (but using Simons' spelling), are shown on Map 2.



Map 2: Languages of Malaita

2.1.2 Malaitan languages brought to Fiji

As mentioned above, it is difficult to determine the numbers of speakers of the different languages brought to Fiji from a large island when it is not known exactly where on the island the labourers came from. However, there is some information available on more precise origins of at least some of the Malaita labourers. It comes from the contract lists and lists of returned labourers for a recruiting voyage of the *Rotuma* in 1899 and two voyages of the *Clansman* in 1908 and 1910, found in the journals of the government agents.⁸ These lists include the exact places where the labourers were recruited or returned. Matching the place with the language spoken in the area gives us some idea of the languages the labourers spoke. This information is summarised in Table 2-2.

According to this information, all the languages were probably represented with the exception of the most minor, Oroha. The North Malaitan dialects had the most speakers, especially Lau.⁹ The next largest group was Kwaio. These figures approximate the proportions of descendants of labourers from each language area found in Fiji today.

Table 2-2: Languages of some Malaitan labourers, 1899-1910

PLACE	RECRUITS	RETURNS	LANGUAGE/DIALECT	TOTAL
Sio Bay	19	20	NM/Lau	
Urası Cove	48	23	NM/Lau	
Ataa Cove	16	1	NM/Lau	
Port Adam	11	--	NM/Lau	163
Cape Astrolabe	20	--	NM/To'abaita	
Bitā'ama Harbour	--	1	NM/To'abaita	
Coleridge Bay	--	4	NM/Baegu, Fataleka	
Kwai Bay	2	1	Kwai	3
Fiu Bay	--	1	Kwara'ae	
Auki	9	1	Kwara'ae	11
Alite Bay	5	14	Langalanga	19
Bina	--	1	Kwaio	
'Olomburi	31	1	Kwaio	73
Uru Bay	37	4	Kwaio	
Su'u Bay	13	4	Dori'o	
Baunani	1	--	Dori'o	18
Takataka Harbour	27	--	'Are'are	27
Sa'a	3	--	South Malaita	
Su'upeine	4	--	South Malaita	14
Mapo	8	--	South Malaita	
UNKNOWN	1	2		3

2.2 Use of imported languages

In Fiji the plantation language was Fijian or Pidgin Fijian (Siegel 1982), generally used by Europeans and Fijians to communicate with the imported Pacific islands labourers and by the linguistically diverse labourers among themselves as a lingua franca. However, there is evidence that many of them continued to speak their home languages within their own groups in Fiji. In fact, some labourers were able to survive without learning any kind of Fijian or English, as evidenced by court records which often mention one "unable to speak any but his own language" (e.g. CSO 1936/1883). This was possible because of the policy on the plantations of allowing labourers speaking the same language to live and work together (Forbes 1875:61-62; Gordon-Cumming 1882:333).

In urban areas also, the home languages were maintained to some extent. The following observation was made by Rev. R.H. Codrington during a visit to Fiji in 1893: "The great mass of the Melanesian Christians in Suva come from Malauta [sic] in the Solomon Islands; among themselves they talk their own tongue, in their intercourse with others they speak the current Fijian of the place ..."¹⁰

But although the settlers may have continued to use their own languages among themselves, they generally did not teach them to their children born in Fiji. Therefore, the children, who mostly had Fijian mothers, learned only Fijian (Kuva [1974]:14). According to informants, even when both parents were from the same island and spoke the same language, they still spoke to their children in Fijian. Thus, it was reported for one community in 1941: "The bulk of the older people can still speak their mother tongue ... but the younger people are to all intents and purposes Fijians".¹¹

It also appears that the older people stopped using their mother tongues after they had lived in Fiji for many years. As one report points out, "Fijian is their adopted tongue".¹² Of nine ex-labourers interviewed by Kuva ([1974]: 13) in the early 1970s:

Five said that they still knew and used their languages, two lost some of the vocabulary, one that he knew only a bit, and one said that he knew less of his own language than Fijian. However, all used Fijian daily in their homes.

Nearly all of the men and women I interviewed speak standard colloquial Fijian as their main language, and are indistinguishable from Fijians in their speech. Only Jioji Abunio (section 1.3) spoke Pidgin Fijian. A few of the first generation also know a little English, but not Melanesian Pidgin. One man, Pita Teqe, also speaks fluent Fiji Hindustani indistinguishable from that of the Fiji Indians.

The only ones who still speak their ancestral language in daily life are some of the descendants of Kiribati labourers who live in communities, such as Veisari, which include free immigrants from Kiribati. The descendants of labourers from Vanuatu have no memory of their parents speaking in their native languages, and none of them can speak even a word of any Vanuatu language. As one informant, Rosalia Mataro, says, "When our parents died, their language also died." The descendants of those from the Solomons, however, all remember the previous generation speaking to each other in their Solomons languages and some still know how to speak these languages.

2.3 Solomons languages in Fiji today

The descendants of labourers from the Solomon Islands differentiate the main languages of their forefathers according to the names of their islands as they are known in Fiji: Malaita, Kalekana (Guadalcanal), Bugotu (Isabel), and Makila (Makira). All the *Kai Solomone* I interviewed trace their ancestry back to Malaita. The names they use for some of the language areas of Malaita are Langalanga, Marato, Koio, Vataleka, Bali, and Wai. Interviews revealed that Marata refers to 'Are'are (it is the 'Are'are word for Malaita), Koio is Kwaio,¹³ Vataleka is Fataleka and Kwara'ae, and Bali is To'abaita (actually the north-west coastal part of the To'abaita area). According to informants, the name "Wai" did not exist on Malaita. There in the local language the people were called *tō i asi* *people of the sea* or *coast dwellers*. This translated into Fijian as *kai wai*. In modern times, their language has become known as Lau.

The five villages in the Wailoku settlement are named after some of these Malaita language groups (Kuva [1974]:24): Wai, Marata, Koio, Vataleka, and Balibuka (supposedly Bali combined with Buka, referring to New Guinea islanders). It should be pointed out that only a few of the first generation and none of later generations know about the connection of the Fiji names with the language groups on Malaita.

In the past, the general pattern in the settlements was that each language-culture group used their own language in speaking to other groups if they did not use Fijian. Similar language use in intergroup contact in North Malaita has been described by Ross (1973:50): "When speakers of different dialects interact, each for convenience uses his own dialect but understands the other's." This

"passive bidialectalism" also occurred in the past, as described by Ivens (1930: 28-29), who also mentions linguistic accommodation as well as maintenance of linguistic boundaries:

When the men or, as happens at market, the women of two different peoples forgather, each person uses his or her own language, and the listener in each case understands ... many individual words of another's speech are known through practice, and a man, when speaking to another, will often make the necessary consonantal changes in his own words in order to agree with the practice of the other. However, in the main, each man talks his own language, and indeed people who dwell among those of another speech seldom seem to learn that speech, but continue to use their own tongue.

But the Fiji informants also say that Wai was used to some extent as a lingua franca in the settlements. (One compared it to Bauan.) In Wailoku it was reportedly spoken to some extent by people from the Vataleka and Koio villages. Many of the informants say that Wai was also used as a lingua franca back on Malaita.¹⁴ The Kai Wai were fishermen and traded fish for garden produce with other language groups. The Wai language was used for this trade, they say, because it was easy to understand and learn.

Only a very few speakers of Malaitan languages are left in Fiji. Some informants from the Koio group know a few words of their fathers' language, but none are fluent, although they say there are some fluent speakers left alive. There are also supposedly one or two speakers of Kwara'ae still living. But most of the first generation informants from the Wai group know at least a few words of the language, and a few know it well. These informants and the Wai language are described in the following section.

3. DESCRIPTION OF WAI

In the following description, Wai is compared with available information on the existing North Malaitan dialects. Word lists for the different varieties of North Malaitan are available in Tryon and Hackman 1983, but the only dialects which have been studied in detail are Lau and To'abaita. Short grammars of two varieties of Lau were written by Ivens (1921, 1929), and a dictionary was compiled by Fox (1974). A dictionary of To'abaita was written by Waterson in 1924 (referred to in Simons 1982). Lichtenberk's (1984) study of a subdialect of To'abaita (called To'aba'ita) is the most detailed description of a North Malaitan dialect. It may be relevant to other dialects, however, as according to Ross (1973:50), "all North Malaitan dialects share essentially the same grammar".

3.1 Informants

Although Wai is no longer used for everyday communication, six informants knew it well enough to record an extended discourse. Their backgrounds and how they learned the language are given here.

The first is Jone Gagalia of Wailoku. Both his parents were from Manaoba Island off Malaita, a Lau speaking area, and in contrast to others they generally spoke to him in their own language, which Jone says was Wai. He is by far the

most fluent of the informants. Jone's half brother, Pita Tege of Waidradra, says he was born in the Solomons and came to Fiji as a child with his mother, but he is nowhere near Jone in fluency.

Charlie Kelo, of Naviavia, learned Wai by listening to the old people talk because his parents spoke to him only in Fijian. His father was from the Fataleka area, and his mother was a part-European from Fiji. His wife, Eni Birena, learned Wai by listening to the conversation of her parents, who also spoke to her in Fijian even though they were both from the Wai area of Malaita.

Makitalena of Waidradra learned a bit of Wai from her father, who tried to teach her his language (her mother was Fijian), but he died when she was very young. Jone Mawia of Wailoku also learned Wai from his father, who came from the Suafa area (which he says is Wai). He relates that his father would slap him if he didn't speak it properly.

Tape recordings of the Wai informants were transcribed and analysed with the help of several informants from the different dialect areas of North Malaita.¹⁵ The general opinion of these informants is that Wai contains a mixture of vocabulary from the various NM dialects, especially Lau, To'abaita, Baelelea, and Baegu. They also mention the influence of Fijian. Those from Lau think that basically Wai sounds like the language spoken by inlanders (that is, the other NM dialects). However, those from non-Lau areas think the rhythm and intonation sound like Lau. They also think that some of the Wai speakers are not fluent and sound like they are just learning the language. This is not only because of hesitation, they say, but also because of leaving out certain words and speaking in what they say is a simple way. These two characteristics, mixture of dialectal features and relative formal simplicity, in addition to the observable influence of Fijian, make Wai different from any dialect currently spoken on Malaita.

3.2 Wai phonology

The phonology of Wai (see Table 3-1) is not identical with that of any particular dialect of North Malaitan (NM) and appears to be influenced by Fijian. But, for the most part, NM and Fijian phonologies (including dialects other than Bauan) are similar, and the overlapping areas are found in Wai. The vowels of Wai are the same as those of both NM and Fijian, including phonemic vowel length. The consonants also appear to be basically those found in both. Wai includes the labiovelar [kw], [ŋgw] (Fijian qw), and [ŋw] (Fijian gw), found not in Bauan but in Western Fijian and some eastern Vitilevu dialects (Geraghty 1983:42-47) as well as in some NM dialects. Also, some Wai speakers use /h/, also not in Bauan but in some Western Fijian dialects and in Lau (Malaita). The following Fijian consonants not found in NM are also not in Wai: c [č] and dr [ɟʀ].

The main differences between Wai and NM phonology appear to be the result of transfer (both positive and negative) from Fijian. First, NM /f/ (To'abaita [ɸ]) is realised as /v/ [β] for most speakers: e.g. vera (Lau fera) *village*, kavu (Lau kavo) *water*. Note, however, that Ivens (1929:324) mentions /f/ sometimes going to /v/ on Sulu Vou island where he studied Lau.

Second, voiced stops are usually prenasalised in Wai as in Fijian. In Lau, prenasalisation either does not occur or it is very slight. In To'abaita, it varies from strong in word-medial position to weak or nil in word-initial position (Lichtenberk 1984:3). Ivens (1929:323) notes, however, that [ŋg], [nd], and [mb] do not appear in Lau but that they are used in the "hill languages".

Third, some phonemes absent in Standard Fijian but present in NM are also absent in Wai. The NM glottal stop, especially in initial position, does not occur in the speech of most Wai speakers: e.g. ave (Lau 'afe) *wife*, iya (Lau T'a) *fish*. The voiceless interdental fricative /θ/ in To'abaita and Baelelea also does not occur in Wai. Instead, the Lau, Baegu, and Fataleka reflex /s/ or /h/ occurs.

Table 3-1: Wai phonology and orthography

CONSONANTS	bi- labial	labio- dental	apico- dental	apico- alveolar	alveo- palatal	dorso- velar	labio- velar	glot- tal
STOPS								
voiceless	p		t			k	kw	
voiced	(b)		(d)			(g)	(gw)	
prenasalised	mb		nd			ŋg	ŋgw	
NASALS	m		n			ŋ	ŋw	
FLAP/TRILL				r				
FRICATIVES								
voiceless	(f)	[ɸv̥f]	(θ)	s				h
voiced	v	[β]						
LATERAL				l				
GLIDES	w				y			
Note: Phonemes in parentheses are found in some NM dialects but not in Wai.								
VOWELS		front	central		back			
HIGH		i			u			
MID		e		o				
LOW			a					
Phonemic vowel length is marked 'ˉ', e.g. ā, ō.								

3.3 Lexicon and pronoun systems

3.3.1 Mixture of marked forms

Apparent dialect mixing in Wai can be observed mainly in the lexicon and in the pronoun system in the concurrent use of some marked NM lexical items. By marked items, I mean those that in North Malaita would identify the speaker as belonging to a particular language-culture group: either one or more of the main dialect areas or coastal versus bush. Other authors have made the distinction between the language spoken by the "coastal people" - those living on the small off-shore islands¹⁶ - and that spoken by the "hill" or "bush" people - those living inland. The coastal language is Lau, and the bush language comprises the other NM dialects. For example, Ivens (1929:323) points out that the language spoken at Ataa Cove (at the southern end of Lau lagoon) "has a closer affinity to the languages of the hill peoples of the mainland than has Lau proper". Fox (1974) also gives some "hill words" in his dictionary of Lau, such as maleu *sleep*.

The names Baelelea, Baegu, and Fataleka themselves may be comprised of such marked items. For example, according to Ross (1973:50), the name Baelelea supposedly comes from its speakers' marked habit of reduplicating the word *lea go* to *lelea* (*bae* is *speak* or *say* in Lau and Baegu). The name Baegu is from its speakers' way of answering the greeting "Where are you going?" with *nao gu nothing*, implying "I do not wish to say". And the name Fataleka is composed of the descriptive use of two marked Fataleka forms: *fata speak, say* and *leka go* (Ivens 1930:24; Ross 1973:50).

Marked items may also include words that are known by a group but cannot be used by them because of word tabooing, and therefore, these items are marked as belonging to other groups. Word tabooing in Malaitan languages has been described in detail by Keesing and Fifi'i (1969) and Simons (1982). It is taboo to say the name of a dead ancestor and to use any common words which are components of the name. Tabooed words are replaced by new forms using a variety of methods (Keesing and Fifi'i 1969:166-168) including borrowing, semantic shift, phonological modification, or simply adopting an already available alternative form. In addition to the forms resulting from word tabooing, some languages have alternative honorific forms and forms used to refer to women (Simons 1982:218n). Therefore, speakers of Malaitan languages have passive knowledge of several synonymous forms for many lexical items, even though some may not be actively used for cultural reasons.

In Wai, sometimes only one marked form appears. For example, NM informants noted that the demonstrative *this* is *nē* or *ne'e* in bush areas and *na* in coastal areas (Lau). Another example is bush *kūfia drink (it)* versus coastal *gwoufia*. In Wai only the bush alternative is found for these items:

- (1) nau tō i vera nē
1S stay LOC village this
I live in this village.
- (2) arai nē kū-vi-a kwakwanga uri-a na koito
old.man this drink-TR-OM kava like-OM DEF dog
This old man drinks kava like a dog.

On the other hand, some marked coastal forms also exist in Wai, such as items with *h* like *haitamana know*.

There are also examples in Wai of items marking a particular "bush language" group, such as BGU *andea make, do* and the proper article *sa*, and also the TOB intensifier *bo'o*:¹⁷

- (3) o ande-a tā
2S do-3S what
What will you do?
- (4) hata-na arai sa jek k.
name-3S.POS ratu PRP J. K.
His name was Ratu Jack K. ("Ratu" is explained in section 3.3.2)
- (5) diana mamana bō
good true INT
really good!

Some lexical items in Wai mark two or three groups of NM, excluding others, for example: BLE, BGU, FTK *sēki here*; LAU, BLE, BGU *lea go*; and LAU, TOB *tā what*.

While the examples above represent the consistent use of one of several possible marked items, there are also instances of two items in use concurrently. Some of these are LAU, BLE, BGU bae along with TOB ɲata *talk* and BLE, BGU, FTK ɲgā along with LAU tē *mother*. Most often different speakers use only one of the alternatives as in the following examples (from two different speakers):

- (6) a. ɲgā kamelu mae nā
mother LXP die PFT
Our mother has died.
- b. tē nau nia mae sui nā
mother 1S 3S die COMP PFT
My mother has died.

However there are instances in which two marked items may be used by the same speaker, even in one utterance:

- (7) molu ɲata diana, molu bae
2P talk good 2P talk
You talk well, you talk.

Wai *not*: There is wide variation among speakers in negative constructions. Four speakers use only one of two marked negative verbs (NEGV), either LAU, BLE laŋi or BGU, VTL nao (sometimes shortened to na), usually followed by the negative particle (NEG) si and often without the 3S subject marker e:

- (8) maka kia laŋi si lea lau
father 1IP NEGV NEG go again
Our father didn't go back again.
- (9) kera nao si haitamana
3P NEGV NEG know
They didn't know.

One informant used only the negative accomplished mode subject marker as in To'abaita (see Lichtenberk 1984:7):

- (10) maka kia ke-si riki-a vera kia
father 1IP 3P.SM-NEG see-OM village 1IP
Our fathers didn't see our village.

One informant used both constructions (as in Lau):

- (11) a. malevo laŋi si baita
money NEGV NEG big
The money (pay) isn't a lot.
- b. mi-si voli-a te-si dō diana
LXP-NEG buy-OM one-NEG thing good
We don't buy anything good.

To see if Wai can be attributed to any particular dialect, forms for 50 lexical items which show variation in NM are compared in Table 3-2 below (mainly from the word lists in Tryon and Hackman 1983). Word lists, of course, cannot be expected to provide complete information, especially in light of the degree of synonymy in Malaitan languages because of word tabooing. Nevertheless,

Table 3-2: 50 lexical items in North Malaitan and Wai

	LAU	TOB	BLE	BGU	FTK	WAI
<i>blood</i>	'abu	'ambu	'ambu	'ambu	'ambu	'ambu
<i>bone</i>	ogi/suli	oki	oki	oki	suli	oki
<i>egg</i>	falake	falake	θāluna	farake	falaka	valake
<i>grease</i>	ragufa	raraŋa	raraŋa	raraŋa	ndūla	ranguva
<i>head</i>	kete/gwou	gwau	ŋgwau	ŋgwau	gwau	ŋgou
<i>mouth</i>	faka	foko	faka	ŋindu	ŋindu	vaka
<i>nose</i>	ŋoŋora	ŋgwalusa	ŋoŋora	ŋoŋora	ŋoŋora	ŋoŋora
<i>bird</i>	manu	θa'aro	θa'aro	saro	nō/θārō	manu
<i>dog</i>	kui	kui/kuri	koito	ŋgiri	kui/kukui	kui/koito
<i>leaf</i>	'aba	rea	'amba	'amba	'amba	amba
<i>bread-fruit</i>	baleo/rauai	mbale'o/kekene	kekene/rau'ai	rau'ai/kekene	rau'ai/kekene	rauai
<i>man</i>	wane	ŋwane	ŋwane	ŋwane	ŋwane	(ŋ)wane
<i>woman</i>	geni	kini	kēni	keni	keni	keni/geni
<i>child</i>	wela	ŋwela	ŋwela	ŋwele	ŋwele	(ŋ)wela
<i>father</i>	mā	maka	maka	mā	mā	mā/maka
<i>mother</i>	tē	θaina	ŋgā	ŋgā	ŋgā	tē/ŋgā
<i>name</i>	sata/hata	θata	θata	sata	sata	hata/ata
<i>earth</i>	gano	ŋgano	θaengano	saengano	gano	ŋgano
<i>garden</i>	raku	o'ole	raoa/ole	ole	ole	rawa
<i>meat</i>	mariko	mariko	maroko	fasio	fasi	mariko
<i>stone</i>	fou	fau	fau	fau	fau	vou
<i>village</i>	fera	fanua/toa	fera	fera	ferē	vera
<i>sky</i>	salo	θalo/mamaŋā	lofonaθalo	mamaŋā	maŋa	salo
<i>wind</i>	ōru	θauθau	θauθau	komburu	sasaulē	oru
<i>day</i>	dani	ndani	ndani	sato	asoā	daŋi
<i>this</i>	na	ne/ne'e	nē	nē/neki	rune'e	nē
<i>what</i>	tā	ta	tae	tē	tē	ta
<i>where</i>	fai/fei	fai/fei	fai	fai	fai	vei/fei
<i>who</i>	tei	tei	tai	tī	tī	tei
<i>yes</i>	iuka	iu	iuka	iuka	iuka	io
<i>no</i>	laŋi	'e a'i	laŋi	nao	nao	laŋi/nao
<i>here</i>	sēgi	kune'e	sēki/nē	seki	seki	seki
<i>left</i>	mouli	mauli	moli	moli	maluli	mauli
<i>fear</i>	mou	ma'u	mau	mau	mau	mou
<i>good</i>	diana	le'a	ndiana	diana	lea	diana
<i>cold</i>	gwagwari	ŋgwangwari	ŋgwangwari	ŋgwari	gwari	ŋgwangwari
<i>old man</i>	waro	arai	ŋwaro	ŋwaro/arai	waro	arai
<i>rotten</i>	fura/dura	fura	ndura	ndura	dūla	vura
<i>heavy</i>	gulu	kulua	kulu	kulu	kulu	ŋgulu
<i>play</i>	lalaba	θaθo	θaθo	saso	saketo	lalamba
<i>go</i>	lea	lae	lea	lea	leka	lea
<i>know</i>	haitamana	θaito'omana	θaitamana	saitamana	saiana	saitamana/haitamana
<i>drink</i>	gwou	ku'u	kū	kū	kū	kou/kū
<i>eat</i>	faŋa	faŋa	faŋa	faŋa/ani	ania	vaŋa/ania
<i>laugh</i>	waela	waela	ŋwaela	ŋwaela	ga'a	ŋwaila
<i>sit</i>	gwouru	ono	ŋgoru	ŋgoru	--	ŋgoru
<i>sleep</i>	teo	mbTŋa	maleu	maleu	malīu	maleu/teo
<i>do</i>	ilia	'ufi/ilia	ilia	andea	sisia	andea
<i>speak</i>	bae	ŋata	bae	bae	fata	ŋata/bae
<i>also</i>	laugo	laubo	laubo	laugu	laugo	laugu/laungo

comparing Wai with the different dialects may give us at least some idea of its linguistic affiliations. In making the comparisons, I have first taken into account the regular phonological differences which result from transfer from Fijian, such as *f* to *v* and loss of glottal stop.

The comparison shows that 29 Wai forms also occur in LAU, 21 in TOB, 27 in BLE, 23 in BGU, and 14 in FTK. In some cases only one dialect has forms corresponding with those of Wai, while in other cases all dialects have corresponding forms. Mostly, however, correspondences occur in different combinations of two or more dialects. The frequencies of the different combinations are illustrated in Table 3-3 for the four most important dialects. This table shows a complex pattern, with items attributable to nearly every possible combination of dialects. Lau appears to be the most important, although not by much. However, if the presence of prenasalised voiced stops in Wai is attributed to transfer from Fijian rather than to the bush dialects, then the contribution of Lau is even more significant, with an additional seven items.

Table 3-3: Combinations of occurrence for Wai items in NM

LAU	TOB	BLE	BGU	frequency
+				6
	+			3
		+		1
			+	2
+	+			6
+		+		2
+			+	1
	+	+		1
	+		+	0
		+	+	4
+	+	+		2
+		+	+	8
+	+		+	0
	+	+	+	5
+	+	+	+	4
				5
29	21	27	24	50

The Wai pronouns present a different picture. The independent or free pronouns (by far the most frequently used in Wai) are shown in comparison to NM dialects (based on Simons 1980) in Table 3-4. Only three of the 11 Wai pronouns can be attributed to Lau. In contrast, six occur in TOB, nine in BLE, ten in BGU, and eight in FTK. In fact, none of the Wai independent plural pronouns appear to be derived from Lau. It may be argued that again there is the influence of transfer: since the voiced velar stop *g* does not occur in Fijian without prenasalisation, it may have come into Wai as the voiceless stop *k*. But even if this argument is accepted, that still leaves four out of eight plural pronouns which appear to be derived from dialects other than Lau if the vowels are compared.

Table 3-4: North Malaitan and Wai independent pronouns

	LAU	TOB	BLE	BGU	FTK	WAI
1S	nau	nau	nau	nau	nau	nau
2S	'oe	'oe	'oe	'oe	'oe	oe
3S	nia	nia	nia	nia	nia	nia
1ID	goro	koro	koro	koro	koro	koro
1IT	golu	kulu	kolu	kolu	kulu	kolu
1IP	gia	kia	--	kia	kia	kia
1XD	gemere	kamere'a	kamere	kamiri	karo	
1XT	gemelu	kamili'a	kameli	kamelu/kameli	kalu	kamelu
1XP	gami	kami	--	kani	kani	
2D	gomoro	kamoro'a	kamoro	kamuru	kamoro	kamoro
2T	gomolu	kamulu'a	kamulu	kamolu	kamulu	kamulu/kamolu
2P	gamu	kamu	--	kamu	kamu	
3D	daro	kero'a	keroa	kerua	keroa	kerua
3T	dalū	kilū'a	--	--	--	--
3P	gera	kera	kera	kera	kera	kera

3.3.2 Semantic and functional extension

The influence of Fijian has brought about the innovative use of some NM forms in Wai. First, a semantic extension has occurred for a few NM lexical items to fit Fijian culture. Most common is Wai kwakwanga *kava* (Fijian *yaqona*). According to NM informants there is no *kava* drinking on Malaita, and kwakwanga is a plant with leaves similar to those of the *kava* plant found in Fiji. Another is the Wai use of *arai* NM *married man*, *elder* as an honorific title corresponding to Fijian *rātū* (see example (4) above). Also, NM *mouri* or *mauri to be alive* is used in Wai as a greeting corresponding to Fijian *bula*.

Second, a functional shift has occurred for Wai *mai*. In NM *mai* functions only as a directional marker indicating action towards the speaker or the point of reference (see Lichtenberk 1984:18). In Wai, it takes on the additional role, as in Fijian, of a marker of general or distant location (Geraghty 1976: Schütz 1983:354) as in this example:

- (12) e laŋi tō mai luma
 3S.SM NEG stay LOC house
 (It) is not in the house.

3.3.3 Formal simplicity

Wai appears to be linguistically less complex than any of the NM dialects in derivational and inflectional morphology.¹⁸ Although the degree of simplicity varies between speakers, they most often use only one of several available NM grammatical strategies. For example, four out of six speakers use only the negative verb rather than negative SM pronouns (examples (8) to (11) above). All but one speaker use the independent pronouns rather than the subject marking (SM) pronouns in affirmative sentences in which the subject has a human referent.

The co-occurrence of the independent rather than the SM pronoun with the subject NP, as in the following Wai example, is not acceptable to NM informants (see also example (6b)):

- (13) ave nau nia geni i bugotu
wife 1S 3S *woman* LOC B.
My wife is a Bugotu woman.

Although the basic pattern in NM includes the subject marking pronouns as an obligatory part of the verb phrase, they may be omitted stylistically in some dialects (Simons 1980:7), such as in To'abaita when the referent of the subject can be recovered from the context (Lichtenberk 1984:13). However, NM informants thought that Wai speakers' elimination of subject markers was excessive. For example, in correcting the transcript of one Wai speaker, a NM informant inserted the subject marker ku:

- (14) nau [ku] tō i vera nē i viti
 1S [1S.SM] *stay* LOC *village* *this* LOC *Fiji*
I stay in this village in Fiji.

NM dialects have two classes of possession: alienable and inalienable. Like Fijian, inalienable possession is marked by adding a possessive suffix to the head noun. But unlike Fijian, alienable possession is marked simply by placing the independent pronoun after the head noun (Simons 1980:7). For some kinship terms such as *father*, there are two different items, one alienable, one inalienable, for example, TOB θāma-ku and maka nau *my father* (Lichtenberk 1984:54). Wai speakers consistently use the alienable alternatives. One speaker uses the alienable construction where the inalienable one is required in NM:

- (15) ata nau arai sa jon
name 1S *ratu* PRP J.
My name is John.

Another speaker overgeneralises the use of the third person singular inalienable possessive suffix:

- (16) vinda-li-a amba-na
clap-TR-OM hand-3S.POS
Clap your hands. (Literally: Clap his hands.)

Wai speakers also leave out the locative marker i:

- (17) vera nē faŋa diana uri-a faŋa vera kia
village this food good like-OM food village 1IP
Is food in this village good like in our village?

The following examples show reduced derivational morphology in that the nominaliser -la or -lā is not used in Wai where it would be expected in NM:

- (18) a. ma nau na si haitamana bae[-la] nē
and 1S NEGV NEG *know* *talk[-NOM]* *this*
And I don't know this language.
 b. dami[-la] e laŋi
chew.betelnut[-NOM] 3S.SM NEGV
There's no betelnut chewing.

3.4 Discussion

On the basis of preliminary data, it is difficult to ascribe Wai to any one North Malaitan dialect. It appears to be characterised by a mixture of lexical forms which currently differentiate the different dialects, and by the influence of Fijian, especially in phonology. It is generally less complex than any of the NM dialects, resulting in some constructions which are ungrammatical according to NM speakers.

4. SOCIOLINGUISTIC FACTORS AFFECTING WAI

4.1 Isolation from Malaita

Without more data on the linguistic situation in Malaita during the Fiji labour trade, it is difficult to come to any firm conclusions about the linguistic or sociolinguistic factors which led to the differences between Wai and the modern NM dialects. It is possible that some of the features described above represent the state of one or more NM dialects at the time they were brought to Fiji. As Fiji was cut off from the Solomons at the end of the Fiji labour trade in 1911, linguistic changes could have taken place on Malaita but not in Fiji.

For example, if the word tabooing responsible for some of the marked lexical items in NM occurred in a particular area after 1911, it is highly unlikely that it would be reflected in Fiji. It may be that some marked lexical forms indicating modern NM dialect boundaries were not marked at the time NM speakers came to Fiji.

If this were true, we might be able to discount that mixing of lexical items occurred among the Fiji immigrants. But it still would not explain some of the other characteristics of Wai. Furthermore, there is no reason for the *Kai Solomone* to be different from other immigrant communities in which dialect mixing has been reported.

4.2 Dialect mixing

Two processes have been described in the literature which refer to the result of contact between linguistic subsystems such as regional dialects. One is "dialect levelling", defined by Dillard (1972:300) as "the process of eliminating prominent stereotypable features of difference between dialects". Another is "dialect mixing", defined by Samarin (1971:133) as an "amalgamation" of several regional varieties of the same language, characterised linguistically by the incorporation of features from these varieties. Although some writers equate the two processes, they are different in their possible end results. In levelling, the original dialects in contact remain and become more like one another, but in dialect mixing a new, mixed variety can emerge which is used as a *lingua franca* among speakers of the original varieties.

The term "koine" has been used to label such a resultant variety. It comes from the Greek word *koinē* *common* referring to the language which became the *lingua franca* of the eastern Mediterranean during the Hellenistic period. It was based mainly on the Attic dialect, but contained features of several other regional dialects of Greek, such as Ionic. However, it was less complex than any of the contributing dialects (Thomson 1960:34-36). Since that time, the term "koine" has been applied to many other dialects which became regional *lingua*

francas: for example, the ancestor of modern Arabic dialects (Ferguson 1959:616), Hindi (Hartmann and Stork 1973:121), Standard Yoruba (Bamgboṣe 1966:2). This type is called a "regional koine" (see Siegel 1985c).

More recently, the term has been extended to refer to the result of dialect mixing within immigrant communities: for example, Trinidad Bhojpuri (Mohan 1976, 1978), Guyanese Bhojpuri (Gambhir 1981), Fiji Hindustani (Siegel 1975:136), Italian-American (Haller 1981:184), and Israeli Hebrew (based on different literary dialects) (Blanc 1968). This type is called an "immigrant koine". The mixture and formal simplicity observed in Wai along with its sociolinguistic history suggests at first glance that it may be an example of an immigrant koine.

The term "koineisation" has been recently coined to refer to the process of dialect mixing (Blanc 1968; Samarin 1971). Koineisation usually takes place only under certain social conditions. Proximinal contact between linguistic subsystems will not necessarily bring it about, for example, if the desire still exists to maintain linguistic boundaries, as can be seen in North Malaita. As Dillard (1972:300) points out, what is required is "some new phase of contact, such as in migration". Koineisation then can result from what Domingue (1981:150) calls "the need for unification among speakers of different dialects in a new environment".

Once koineisation begins, there is a developmental continuum for koines analogous to that for pidgins and creoles. The first stage is the "pre-koine". It is similar to a pre-pidgin continuum or jargon in that individual strategies are used to modify one's own language or in trying to speak another's, and a socially accepted norm has not yet emerged. These strategies will often result in output characterised by mixture and relative formal simplicity. When a social norm does emerge as the result of informal standardisation, the next stage is reached: the "stable koine". Use of a stable koine may be eventually extended to other areas besides intergroup communication. For example, it may become a literary or standard language, as the original Greek Koine did. This extension of use is accompanied by linguistic expansion, such as in an increased lexicon or more stylistic options. This is the "expanded koine" stage. Finally, the koine may become the first language of a social group, or a "nativised koine". These stages and the processes which lead to them are illustrated as follows:

Stages in koine development

PROCESS	STAGE
koineisation	pre-koine
stabilisation	stable koine
expansion	expanded koine
nativisation	nativised koine

Wai developed under some of the social conditions in which koineisation typically takes place, and its linguistic features of formal simplicity and mixture are also typical of koines. But the variation among Wai speakers indicates that in general it did not get past the pre-koine stage of development. Certain areas of Wai grammar, however, such as the independent pronoun system, do show that at least some stabilisation had taken place. If the *Kai Solomone* had not adopted the Fijian language and culture, Wai might have been further stabilised and developed into a nativised koine, like the Fiji Hindustani of the Fiji Indians (Siegel 1975, 1983; Moag 1979).

4.2.1 Language attrition

Although koineisation leads to a variety which is less complex than any of the contributing varieties, most writers point out that the formal simplicity resulting from koineisation is much less drastic than that resulting from pidginisation (Nida and Fehderau 1970; Samarin 1971; Mohan 1976). Therefore, koines are never "structurally discontinuous from their linguistic parents" (Gambhir 1981:185). However, the fact that many Wai sentences are unacceptable to NM speakers because of their lack of certain grammatical features indicates that the degree of simplicity in Wai may be too excessive to be the result of only koineisation. Other processes must be considered, and it may be that the Wai spoken by the informants does not necessarily represent the language spoken by their parents.

The first possibility is that this lack of complexity indicates pidginisation itself, the result of incomplete language learning. In fact, the overgeneralised use of the 3S inalienable possessive suffix, the use of the independent rather than SM pronouns, and the lack of locative marking are also characteristics of Pidgin Fijian (Siegel 1982). Only one of the informants learned Wai as his first language along with Fijian; all the others learned it later. (This would explain the phonological transfer from Fijian.) But without adequate opportunity to use the language, it may never have been learned completely. The reason for this restricted use was that Wai was being displaced by Fijian; in other words, it was (and still is) a dying language.

In her studies of language death, or language attrition over successive generations, Dorian (1973, 1982, 1983) describes how parents may speak the dying language to each other but not to their children. She says that in immigrant communities, for example, the utility of the mother tongue is seen to be reduced and, therefore, it is often deliberately not transmitted to the children (1982: 46-47). Children in this situation, if they do learn any of their parents' language, become one type of what Dorian calls "semi-speakers" - imperfect speakers of a dying language. It may be that all but one of the Fiji informants are semi-speakers of Wai because of inadequate input during childhood and few opportunities to use the language later. Some of the linguistic features resulting from the process of language death or language attrition, described by several authors (Dorian 1973, 1981, 1983; Hill 1978; Andersen 1982), are similar to those of Wai, such as reduction and regularisation of morphology.

Another explanation for the degree of simplicity is that the Wai informants may have acquired the language thoroughly as children but attrition occurred from lack of use (Andersen 1982:85) - especially after most of the NM native speakers had died. All the Wai speakers said that they had not used the language for many years, and some said they had trouble remembering some of the words. Also, some of the NM informants had the impression that at least one speaker started off speaking as if he was just learning the language, but later in the discourse sounded more fluent. Examples (14) and (16) above illustrate reduced and regularised morphology in Wai. But the same speaker comes up with the following more acceptable sentences later in the same discourse:

(19) vinda-li-a amba-molu
clap-TR-OM hand-2P.POS
Clap your hands.

(20) nau ku kū-vi-a kwakwanga nē
 1S 1S.SM *drink-TR-OM kava this*
I drank this kava.

Thus, it could be that for this speaker the attrition was only temporary and the language came back to him as he started to use it.

4.3 Summary

Several sociolinguistic factors could have led to the characteristics of Wai as described above. First, emigration probably disrupted cultural traditions, and the need for solidarity in a new environment broke down boundaries between language-culture groups. This resulted in the active use of some of the more widely known lexical items that may have been in the passive repertoire of speakers of some dialects or communalects. Thus, the apparent lexical mixing in the language of the Wai informants may be attributed to the process of koineisation which has been described for other immigrant communities.

The morphological reduction and regularisation, however, are more likely the result of other sociolinguistic processes. These may be individual language attrition from lack of use, or community level language attrition (language death) resulting in incomplete language learning in individuals.

ABBREVIATIONS

ART	article	NOM	nominaliser
BGU	Baegu	OM	object marker
BLE	Baelelea	P	plural
COMP	completive	PERF	perfective
DIR	directional	POS	possessive marker
FTK	Fataleka	S	singular
I	inclusive	SM	subject marker
INT	intensifier	T	trial or paucal
LOC	locative	TOB	To'abaita
NEG	negative particle	TR	transitive marker
NEGV	negative verb	X	exclusive
NM	North Malaitan		

NOTES

1. This paper came out of a larger study, "Plantation languages in Fiji", sponsored by the Department of Linguistics, Research School of Pacific Studies, The Australian National University. Some of the linguistic data were collected with the help of Darrell Tryon. I would like to thank the following people for their assistance in the field: Tevita Nawadra, Paul Geraghty, the staff of the National Archives of Fiji, the Tutu community, Fr Luke Oli, Fr Ilai Lakavutu, and Aduru Kuva. Also thanks for comments on an earlier draft to Walter Seiler, Roger Keesing, Don Laycock, Lois Carrington, Christine Jourdan, and Frank Lichtenberk.

2. Report on Polynesian Immigration for 1884.
3. References to the official correspondence of the Fiji Colonial Secretary's Office are abbreviated CSO, followed by the minute paper number/year.
4. *Church Gazette* (Diocese of Polynesia) no.11 (1927:8).
5. Letter from W.E. Moren of Auckland, New Zealand to Mr Long, August 1946 (Church of England records, National Archives of Fiji).
6. *Anglican Church Gazette: Golden Jubilee issue* (1958):27.
7. *Anglican Church Gazette: Golden Jubilee issue* (1958):27.
8. Journals number 64, 65, and 66, F. Otway and W.R. Bell, National Archives of Fiji.
9. The figures can only be approximate because after the mid 1880s, most recruits were 'bush people' who spoke dialects other than Lau. These people came down to the coast (i.e. to Lau-speaking areas) in order to sign up when a recruiting ship arrived (Corris 1973:32-36).
10. *Anglican Church Gazette* (October 1893):161.
11. *Church Gazette* no.66 (April 1941):10.
12. *Church Gazette* no.1 (Diocese of Polynesia), November 1924:13.
13. Roger Keesing (personal communication, 19/7/84) points out that "Koio" is the North Malaitan rendering of Kwaio.
14. This has been verified by Pierre Maranda (personal communication, 17/5/84), who says that at least Kwara'ae people have used Lau. However, Roger Keesing and Christine Jourdan disagree (personal communication, 19/7/84).
15. In September 1982 I worked with Barnabas Lauia, a Lau speaker, and in July 1983 with Manuel Maesua, a Baelelea speaker. Both were then students at the University of the South Pacific. In November 1982 Darrell Tryon and I went to the Marist Brothers Training Centre at Tutu, Taveuni, where we worked with Brother Paul Maefiti (To'abaita), Brother Peter Sukuomea (Baegu), and Brother Timothy Beliga (Baegu), who were all also familiar with Lau. Also, in June 1983, I worked with three North Malaitan students at the Papua New Guinea University of Technology in Lae: Bentley D. Collin (Lau), Enaly Fifira (To'abaita), and Frank Lobo (Baegu).
16. Only since European contact have villages been established on the coast of the mainland. Before, for strategic reasons, they were either on small natural or man-made islands or inland.
17. Abbreviations used here are as in Table 2-1: BGU (Baegu), BLE (Baelelea), FTK (Fataleka), and TOB (To'abaita).
18. The notions of complexity/simplicity in languages are not well understood. I use the terms in a quantitative rather than a qualitative sense; that is, an aspect of one variety is less complex than that of another if it has fewer components and fewer rules, or if it is more consistent. It is not necessarily easier to process psychologically.

REFERENCES

ANDERSEN, Roger W.

- 1982 Determining the linguistic attributes of language attrition. In Lambert and Freed, eds 1982:83-118.

ANGLICAN CHURCH OF FIJI

- 1924- *The Church Gazette*. Suva (published prior to 1924).

BAMGBOSE, Ayọ

- 1966 *A grammar of Yoruba*. West African Monograph Series No.5. London: Cambridge University Press.

BLANC, Haim

- 1968 The Israeli Koine as an emergent national standard. In Joshua Fishman, Charles A. Ferguson and Jyotirindra Das Gupta, eds *Language problems of developing nations*, 237-251. New York: John Wiley.

CORRIS, Peter

- 1973 *Passage, port and plantation: history of Solomon Islands labour migration 1870-1914*. Melbourne: Melbourne University Press.

DILLARD, J.L.

- 1972 *Black English: its history and usage in the United States*. New York: Random House.

DOMINGUE, Nicole

- 1981 Internal change in a transplanted language. *Studies in the Linguistic Sciences* 4/2:151-159.

DORIAN, Nancy C.

- 1973 Grammatical change in a dying language. *Language* 49/2:413-438.
- 1981 *Language death: the life cycle of a Scottish Gaelic dialect*. Philadelphia: University of Pennsylvania Press.
- 1982 Language loss and maintenance in language contact situations. In Lambert and Freed, eds 1982:44-59.
- 1983 Natural second language acquisition from the perspective of the study of language death. In Roger W. Andersen, ed. *Pidginization and creolization as language acquisition*, 158-167. Rowley, Mass.: Newbury House.

FERGUSON, Charles A.

- 1959 The Arabic Koine. *Language* 35/4:616-630.
- 1981 Foreigner talk as the name of a simplified register. *International journal of the sociology of language* 28:9-18.

FIJI: Legislative Council

- 1886 Report on Polynesian Immigration for 1884. Fiji Legislative Council Paper no.23 of 1886.

FORBES, Litton

- 1875 *Two years in Fiji*. London: Longmans, Green.

FOX, Charles E.

- 1974 *Lau dictionary, with English index.* PL, C-25.

GAMBHIR, Surendra Kumar

- 1981 *The East India speech community in Guyana: a sociolinguistic study with special reference to koine formation.* Ph.D. dissertation, University of Pennsylvania.

G

- n.d. *General Register of Polynesian Labourers Introduced to Fiji (three volumes).* National Archives of Fiji.

Geraghty, Paul

- 1976 *Fijian prepositions.* *Journal of the Polynesian Society* 85/4:507-520.
- 1983 *The history of the Fijian languages.* *Oceanic Linguistics Special Publication* 19. Honolulu: University of Hawaii Press.

GORDON-CUMMING, C.F.

- 1882 *At home in Fiji.* New York: A.C. Armstrong.

HALLER, Hermann W.

- 1981 *Between Standard Italian and creole: an interim report.* *Word* 32/3: 181-192.

HARTMANN, R.R.I.C. and F.C. STORK

- 1973 *Dictionary of language and linguistics.* London: Applied Science Publishers.

HILL, Jane H.

- 1978 *Language death, language contact, and language evolution.* In William C. McCormack and Stephen A. Wurm, eds *Approaches to language: anthropological issues*, 45-78. The Hague: Mouton.

IVENS, Walter G.

- 1921 *Grammar and vocabulary of the Lau language, Solomon Islands.* Washington: Carnegie Institution.
- 1929 *A grammar of the Lau language, northeast coast of Big Mala, Solomon Islands.* *Bulletin of the School of Oriental and African Studies* 5: 323-343.
- 1930 *The island builders of the Pacific.* London: Seeley, Service.

KEESING, Roger M. and Jonathan FIFI'I

- 1969 *Kwaio word tabooing in its cultural context.* *Journal of the Polynesian Society* 78/2:154-177.

KUVA, Aduru

- [1974] *The Solomons community in Fiji.* Suva: South Pacific Social Sciences Association.

LAMBERT, Richard D. and Barbara F. FREED, eds

- 1982 *The loss of language skills.* Rowley, Mass.: Newbury House.

LICHTENBERK, Frantisek

- 1984 *To'aba'ita language of Malaita, Solomon Islands. Working papers in anthropology, archaeology, linguistics, Maori studies no.65.*
Auckland: University of Auckland.

MOAG, Rodney F.

- 1979 The linguistic adaptations of the Fiji Indians. In Vijay Mishra, ed. *Rama's banishment*, 112-138. London: Heinemann Educational Books.

MOHAN, Peggy

- 1976 To coin a koine theory. Paper for the Conference on New Directions in Creole Studies, University of Guyana.
1978 Trinidad Bhojpuri: a morphological study. Ph.D. dissertation, University of Michigan.

MÜHLHÄUSLER, Peter

- 1974 *Pidginization and simplification of language. PL, B-26.*
forth- Patterns of contact, mixture, creation and nativization: their
coming contributions to a general theory of language. In C.-J. Bailey and R. Harris, eds *Developmental mechanisms of language*. Oxford: Pergamon Press.

MURDOCK, George Peter

- 1964 Genetic classification of the Austronesian languages: a key to Oceanic culture history. *Ethnology* 3:117-126.

NIDA, Eugene A. and Harold W. FEHDERAU

- 1970 Indigenous pidgins and koinés. *International Journal of American Linguistics* 32/2:146-155.

PARNABY, Owen W.

- 1956 The regulation of indentured labour to Fiji. *Journal of the Polynesian Society* 65/1:55-65.
1964 *Britain and the labour trade in the Southwest Pacific*. Durham, N.C.: Duke University Press.

PAWLEY, Andrew and Timoci SAYABA

- 1971 Fiji dialect divisions: eastern and western Fijian. *Journal of the Polynesian Society* 80:405-448.

RABUKAWAQA, Josua

- 1968 Survey of settlements in the Suva district (interim report). National Archives of Fiji.

ROSS, Harold M.

- 1973 *Baegu: social and ecological organization in Malaita, Solomon Islands. Illinois Studies in Anthropology no.8.* Urbana: University of Illinois Press.

RUSSELL, T.

- 1950 The Fataleka of Malaita. *Oceania* 21/1:1-13.

SAMARIN, William J.

- 1971 Salient and substantive pidginization. In Dell Hymes, ed. *Pidginization and creolization of languages*, 117-140. Cambridge: Cambridge University Press.

SCARR, Deryck

- 1967 *Fragments of empire; a history of the Western Pacific High Commission 1877-1914*. Canberra: Australian National University Press.
- 1973 Recruits and recruiters: a portrait of the labour trade. In J.W. Davidson and Deryck Scarr, eds *Pacific island portraits*, 225-251. Canberra: Australian National University Press.

SCHÜTZ, Albert J.

- 1983 The Fijian language. (draft) University of Hawaii.

SIEGEL, Jeff

- 1975 Fiji Hindustani. *University of Hawaii Working Papers in Linguistics* 7/3:127-144.
- 1982 Plantation Pidgin Fijian. *Oceanic Linguistics* 21/1-2:1-72.
- 1983 Koineization and the development of Fiji Hindustani. MS. Australian National University.
- 1985a Plantation languages in Fiji. Ph.D. thesis, Australian National University, Canberra.
- 1985b Origins of Pacific Islanders in Fiji. *Journal of Pacific History* 20:32-42.
- 1985c Koines and koineization. *Language in Society* 14/3.

SIMONS, Gary F.

- 1980 Morphological reconstruction and change in the pronoun systems of the Malaitan languages. Paper presented at the 3rd New Zealand Linguistics Conference, Auckland.
- 1982 Word taboo and comparative Austronesian linguistics. In Amran Halim, Lois Carrington, and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, vol.3: *Accent on variety*. PL, C-76:157-226.

THOMSON, George

- 1960 *The Greek language*. Cambridge: W. Heffer.

TRYON, D.T.

- 1976 *New Hebrides languages: an internal classification*. PL, C-50.

TRYON, D.T. and B.D. HACKMAN

- 1983 *Solomon Islands languages: an internal classification*. PL, C-72.

WURM, Stephen A. and Shirô HATTORI

- 1981 *Linguistic atlas of the Pacific area, part 1: New Guinea area, Oceania, Australia*. Canberra: The Australian Academy of the Humanities in collaboration with the Japan Academy; PL, C-66.

