

PACIFIC LINGUISTICS

Series C - 117

**CURRENTS IN PACIFIC LINGUISTICS:**  
**PAPERS ON AUSTRONESIAN LANGUAGES AND**  
**ETHNOLINGUISTICS IN HONOUR OF**  
**GEORGE W. GRACE**

edited by

Robert Blust



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

**FOUNDING EDITOR:** S.A. Wurm

**EDITORIAL BOARD:** K.A. Adelaar, T.E. Dutton, A.K. Pawley, M.D. Ross, D.T. Tryon

**EDITORIAL ADVISERS:**

B.W. Bender  
University of Hawaii  
David Bradley  
La Trobe University  
Michael G. Clyne  
Monash University  
S.H. Elbert  
University of Hawaii  
K.J. Franklin  
Summer Institute of Linguistics  
W.W. Glover  
Summer Institute of Linguistics  
G.W. Grace  
University of Hawaii  
M.A.K. Halliday  
University of Sydney  
E. Haugen  
Harvard University  
A. Healey  
Summer Institute of Linguistics  
L.A. Hercus  
Australian National University  
John Lynch  
University of Papua New Guinea

K.A. McElhanon  
Summer Institute of Linguistics  
H.P. McKaughan  
University of Hawaii  
P. Mühlhäusler  
Bond University  
G.N. O'Grady  
University of Victoria, B.C.  
K.L. Pike  
Summer Institute of Linguistics  
E.C. Polomé  
University of Texas  
Gillian Sankoff  
University of Pennsylvania  
W.A.L. Stokhof  
University of Leiden  
B.K. T'sou  
City Polytechnic of Hong Kong  
E.M. Uhlenbeck  
University of Leiden  
J.W.M. Verhaar  
The Hague, The Netherlands  
C.L. Voorhoeve  
University of Leiden

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

PACIFIC LINGUISTICS  
Department of Linguistics  
Research School of Pacific Studies  
The Australian National University  
G.P.O. Box 4, Canberra, A.C.T. 2601  
Australia

Copyright © The Authors  
First Published 1991  
Typeset by Jeanette Coombes and Anne Rees  
Maps for papers by F. Ozanne-Rivierre and M.D. Ross drawn by Theo Baumann

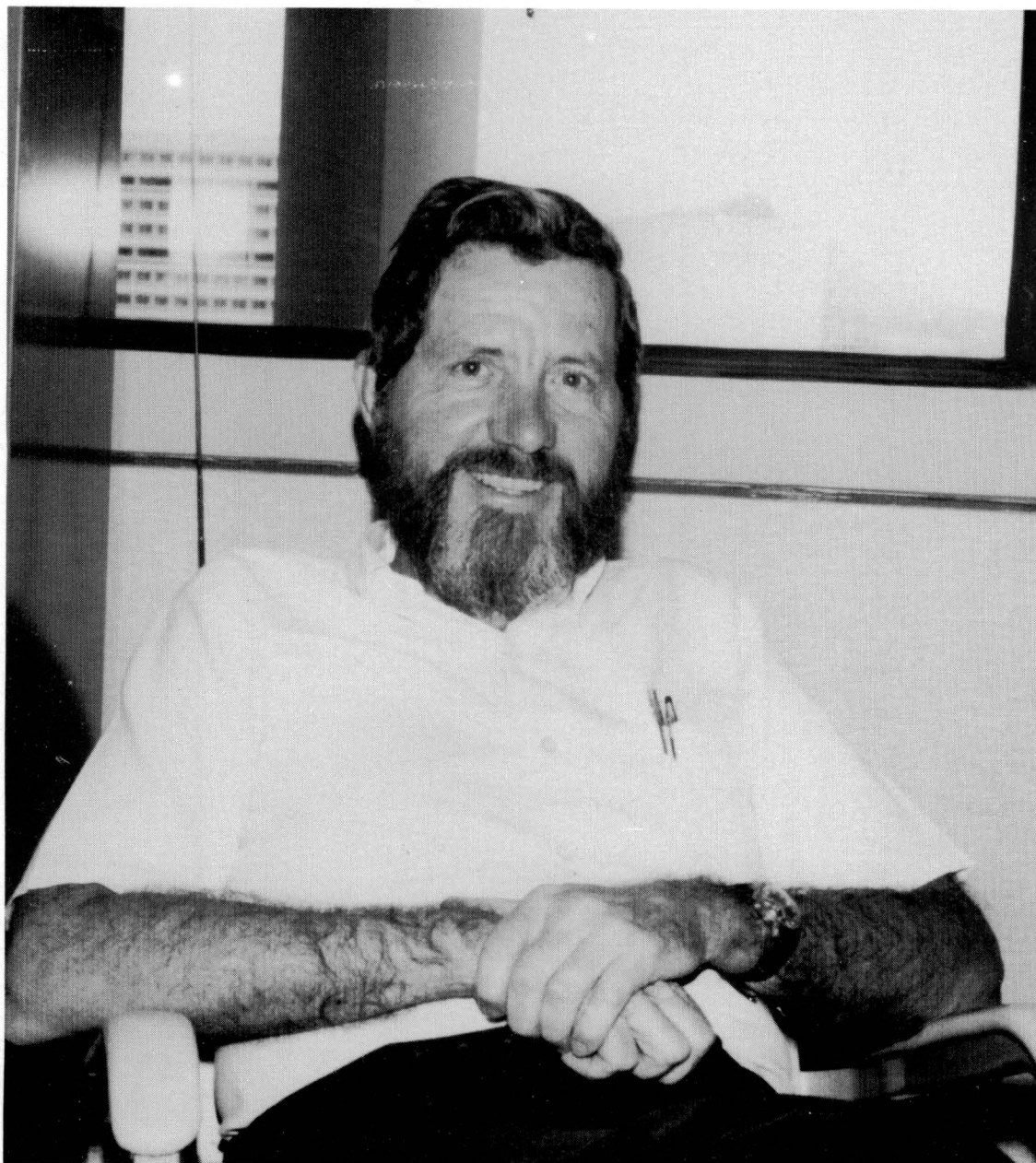
Printed by A.N.U. Printing Service  
Bound by F & M Perfect Bookbinding

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

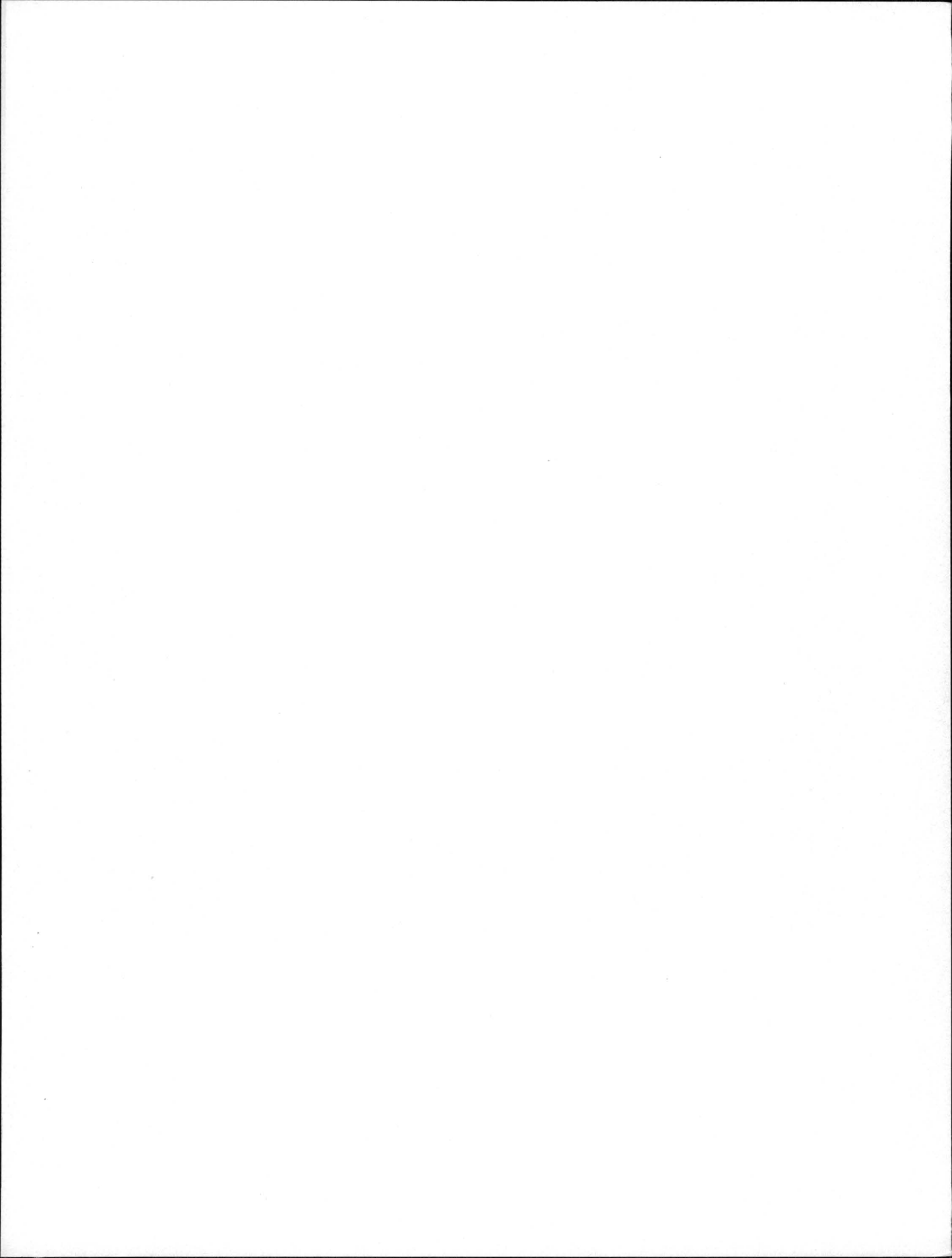
ISSN 0078-7558

ISBN 0 85883 404 9





GEORGE GRACE

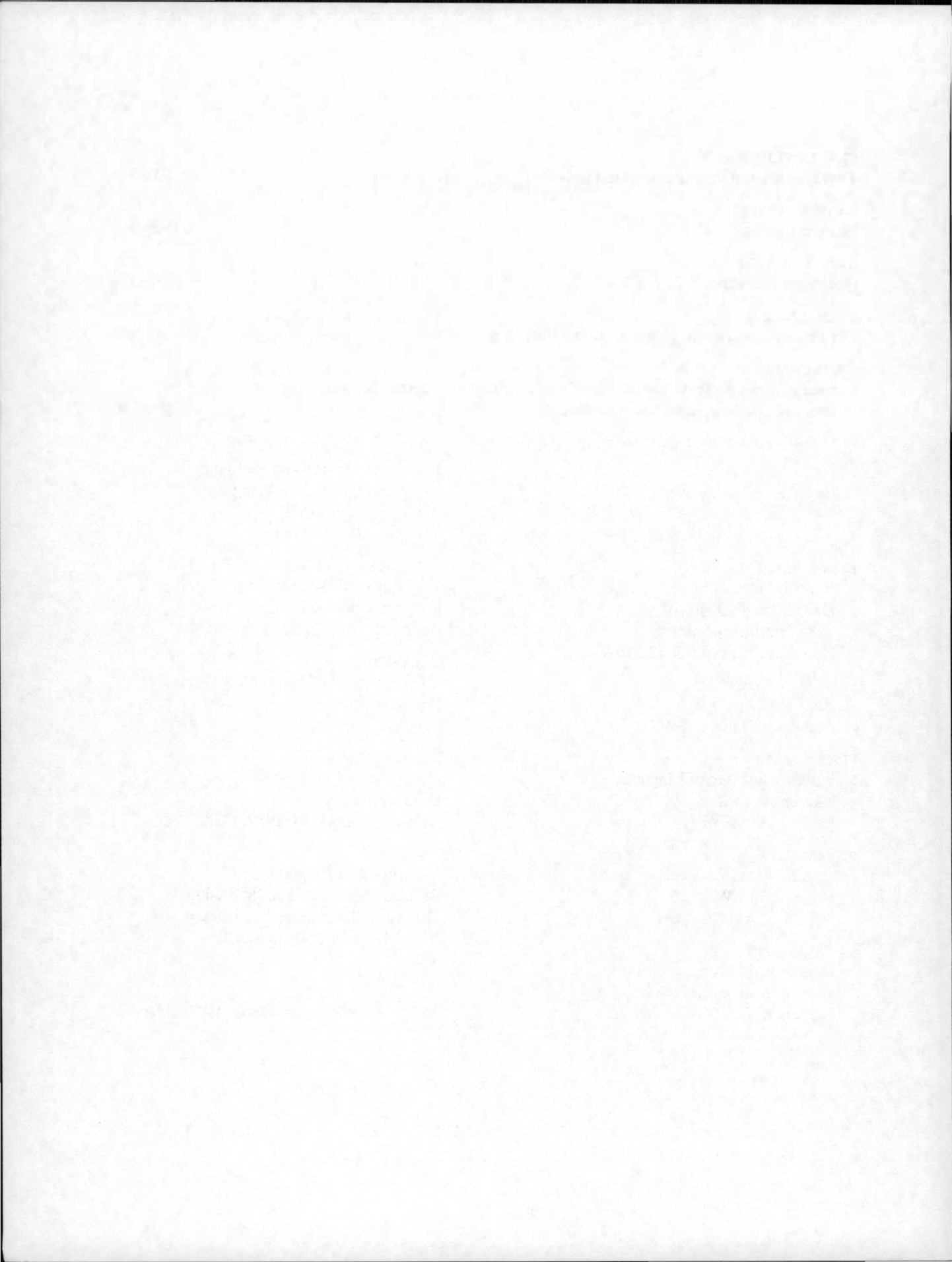


## TABLE OF CONTENTS

Contributors to this volume	ix-x
Editorial note	xi
ROBERT BLUST George W. Grace: an appreciation	1-4
ALBERT J. SCHÜTZ The publications of George W. Grace	5-10
BYRON W. BENDER On the category distributive	11-26
ROBERT BLUST Sound change and migration distance	27-42
ANN CHOWNING Proto Oceanic culture: the evidence from Melanesia	43-75
KENNETH WILLIAM COOK The search for subject in Samoan	77-98
SAMUEL H. ELBERT Shifting attitudes in Hawai'i 1920s to 1990	99-102
ELAINE M. GOOD Determiner forms in Hote	103-116
RAY HARLOW Consonant dissimilation in Maori	117-128
S.P. HARRISON A reanalysis of Proto Polynesian noun phrase marking	129-144
K.J. HOLLYMAN Personalised and non-personalised possession: final consonants in Kumak and other languages of Far Northern New Caledonia	145-154
PATRICIA A. LEE Some practical factors in the early development of linguistics as a science	155-162
PAUL JEN-KUEI LI Vowel deletion and vowel assimilation in Sediq	163-169

FRANTISEK LICHTENBERK Reciprocals and depatientives in To'aba'ita	171-183
JOHN LYNCH A century of linguistic change in Anejom	185-195
RICHARD MCGINN Pronouns, politeness and hierarchy in Malay	197-221
JEFF MARCK Sixteen Nuclear Micronesian verbs	223-239
ROGER F. MILLS Tanimbar-Kei: an Eastern Indonesian subgroup	241-263
MALCOLM W. MINTZ The fossilised affixes of Bikol	265-291
ULRIKE MOSEL The Samoan construction of reality	293-303
CLAIRE MOYSE-FAURIE Relational morphemes and a transitivity suffix in Xârâcùù (New Caledonia)	305-320
F. OZANNE-RIVIERRE Incorporation of genitive relators in the languages of New Caledonia and the Loyalty Islands	321-338
ANDREW PAWLEY How to talk cricket: on linguistic competence in a subject matter	339-368
SUSAN U. PHILIPS Tongan speech levels: practice and talk about practice in the cultural construction of social hierarchy	369-382
KENNETH L. REHG Final vowel lenition in Micronesian languages: an exploration of the dynamics of drift	383-401
KARL H. RENSCH The language of the Noble Savage: early European perceptions of Tahitian	403-414
J.-C. RIVIERRE Loss of final consonants in the north of New Caledonia	415-432
M.D. ROSS How conservative are sedentary languages? Evidence from western Melanesia	433-451
ALBERT J. SCHÜTZ William Anderson's Hawaiian word list	453-464
RON SCOLLON AND SUZIE SCOLLON Mass and count nouns in Chinese and English: a few further Whorfian considerations	465-475
STANLEY STAROSTA Grammar, perception and reality	477-500

CLAUDE TCHEKHOFF Two symmetrical verbal aspects in Oceania	501-508
DARRELL TRYON Wanem Bislama?	509-519
JOHN W.M. VERHAAR Questions and answers in Tok Pisin	521-534
JOHN U. WOLFF The Proto Austronesian phoneme *t and the grouping of the Austronesian languages	535-549
S.A. WURM Language decay and revivalism: the Äyiwo language of the Reef Islands, Santa Cruz Archipelago, Solomon Islands	551-560



## CONTRIBUTORS TO THIS VOLUME

BYRON W. BENDER  
Department of Linguistics  
University of Hawaii  
Honolulu, Hawaii 96822, USA

ROBERT BLUST  
Department of Linguistics  
University of Hawaii  
Honolulu, Hawaii 96822, USA

ANN CHOWNING  
Department of Anthropology  
Victoria University of Wellington  
Wellington, New Zealand

KENNETH WILLIAM COOK  
English Foundations Program  
Hawaii Pacific University  
1188 Fort Street Mall  
Honolulu, Hawaii 96813, USA

SAMUEL H. ELBERT  
3293 Huelani Dr.  
Honolulu, Hawaii 96822, USA

ELAINE M. GOOD  
Summer Institute of Linguistics  
Ukarumpa via Lae  
Papua New Guinea

RAY HARLOW  
Department of Linguistics  
University of Waikato  
Hamilton, New Zealand

S.P. HARRISON  
Department of Anthropology  
University of Western Australia  
Nedlands, WA 6009, Australia

K.J. HOLLYMAN  
182 Devonport Road  
Tauranga, New Zealand

PATRICIA A. LEE  
Department of Linguistics  
University of Hawaii  
Honolulu, Hawaii 96822, USA

PAUL JEN-KUEI LI  
Institute of History and Philology  
Academia Sinica, Nan-Kang  
Taipei, Taiwan 11529  
The Republic of China

FRANTISEK LICHTENBERK  
Department of Anthropology  
University of Auckland  
Auckland, New Zealand

JOHN LYNCH  
The University of Papua New Guinea  
Port Moresby  
Papua New Guinea

RICHARD MCGINN  
Department of Linguistics  
Ohio University  
Athens, Ohio 45701-2979, USA

JEFF MARCK  
Department of Linguistics  
Research School of Pacific Studies  
The Australian National University  
Canberra, ACT 2601, Australia

ROGER F. MILLS  
PO Box 807  
Saugatuck, Michigan 49453-0807, USA

- MALCOLM W. MINTZ  
Asian Studies Programme  
School of Humanities  
Murdoch University, WA 6150, Australia
- ULRIKE MOSEL  
Department of Linguistics  
The Faculties  
The Australian National University  
Canberra, ACT 2601, Australia
- CLAIRE MOYSE-FAURIE  
LACITO  
Centre National de la Recherche Scientifique  
Paris, France
- F. OZANNE-RIVIERRE  
LACITO  
Centre National de la Recherche Scientifique  
Paris, France
- ANDREW PAWLEY  
Department of Linguistics  
Research School of Pacific Studies  
The Australian National University  
Canberra, ACT 2601, Australia
- SUSAN U. PHILIPS  
Department of Anthropology  
University of Arizona  
Tucson, Arizona 85721, USA
- KENNETH L. REHG  
Department of Linguistics  
University of Hawaii  
Honolulu, Hawaii 96822, USA
- KARL H. RENSCH  
Department of Linguistics  
The Faculties  
The Australian National University  
Canberra, ACT 2601, Australia
- J.-C. RIVIERRE  
LACITO  
Centre National de la Recherche Scientifique  
Paris, France
- M.D. ROSS  
Department of Linguistics  
Research School of Pacific Studies  
The Australian National University  
Canberra, ACT 2601, Australia
- ALBERT J. SCHÜTZ  
Department of Linguistics  
University of Hawaii  
Honolulu, Hawaii 96822, USA
- RON SCOLLON AND SUZIE SCOLLON  
PO Box 1149  
Haines, Alaska 99827, USA
- STANLEY STAROSTA  
Department of Linguistics  
University of Hawaii  
Honolulu, Hawaii 96822, USA
- CLAUDE TCHEKHOFF  
1 rue de Marnes  
92410, Ville d'Avray, France
- DARRELL TRYON  
Department of Linguistics  
Research School of Pacific Studies  
The Australian National University  
Canberra, ACT 2601, Australia
- JOHN W.M. VERHAAR,  
't Hoenstraat 30  
2596, HZ, The Hague, The Netherlands
- JOHN U. WOLFF  
Department of Modern Languages and  
Linguistics  
Cornell University  
Ithaca, New York 14853, USA
- S.A. WURM  
Department of Linguistics  
Research School of Pacific Studies  
The Australian National University  
Canberra, ACT 2601, Australia



## EDITORIAL NOTE

Invitations were initially mailed to over 100 potential contributors to this volume. Following a period of fairly heavy correspondence during which titles and abstracts were obtained, the set of probable contributions was narrowed down to about 40 papers. In some cases authors were unable to meet the necessary deadline, and in a few others contributions which were sent could not be accepted for one reason or another. The result is the present set of 34 papers. I wish to thank all contributors for their patience with my editorial meddling. Much appreciation also goes to Andrew Pawley and the staff of the publications section of the Department of Linguistics, Research School of Pacific Studies at the Australian National University for their dedicated – I might almost say heroic – efforts in the production of this volume under considerable time pressure. Finally, I owe a debt of gratitude to Mike Forman for contributing information regarding George's tennis accomplishments, and to Byron Bender for assistance with what turned out to be an inordinately difficult undertaking: obtaining an appropriate photograph of our honoree.

1900

The following table shows the results of the various experiments conducted during the year 1900. The first column gives the name of the experiment, the second column the date, and the third column the results. The results are given in the form of a table, the columns of which are headed by the names of the various quantities measured. The first column of the table is headed by the name of the quantity, and the second column by the value of the quantity. The results are given in the form of a table, the columns of which are headed by the names of the various quantities measured. The first column of the table is headed by the name of the quantity, and the second column by the value of the quantity.

The following table shows the results of the various experiments conducted during the year 1900. The first column gives the name of the experiment, the second column the date, and the third column the results. The results are given in the form of a table, the columns of which are headed by the names of the various quantities measured. The first column of the table is headed by the name of the quantity, and the second column by the value of the quantity.

## GEORGE W. GRACE: AN APPRECIATION

ROBERT BLUST

When the person we are honouring came to Hawai'i to accept a professorship in 1964, he soon learned that he was not the only George Grace in the islands. Time and again he received bills for goods he had never bought, creditors' notices for loans he had never taken out, and menacing letters of various kinds. Patiently, he returned all of these with the explanation that the clamouring mob was after the wrong man. That other George Grace is now well known to the local population as the inventor and developer of the Grace portable flush toilet – his bills paid and his profits handsome. Unlike George Grace the toilet king, our George Grace toils on in relative anonymity, knowing that even his greatest stroke of genius will never touch the queued-up masses as closely as that simple invention by his more entrepreneurial namesake. In recognition of this fact (and others) we have contrived to pay homage to him with this volume.

George's life and work appear rich in irony and incongruity. Born near the northern border of Mississippi in 1921 and raised on the gulf coast of that state, he received his first university degree in Switzerland at the age of 27. Almost thwarted in his efforts to obtain the PhD by an outside committee member who refused to accept his dissertation, he published the work, which became an instant landmark in the field of Austronesian linguistics (Grace 1959). A linguist of exceptionally broad knowledge, interests, and intellectual scope who has made lasting contributions both to historical linguistics and to the philosophy of language, he has remained almost unknown to the general community of linguists.

Following his return to the United States after receiving his Licence-ès-sciences politiques from the University of Geneva, George accepted a position as Junior Research Anthropologist at the University of California at Berkeley. During the summer of 1951 he carried out fieldwork among the Luiseño Indians of southern California, an experience which formed the basis for his collaboration with Alfred L. Kroeber in preparing the Sparkman grammar of Luiseño. By 1953 he was a Research Associate with the Tri-Institutional Pacific Program (a consortium of resources from Yale University, the University of Hawai'i and the Bernice P. Bishop Museum of Honolulu), and the following year a Research Assistant in Anthropology at Yale University. He spent the period 1955-1956 conducting a linguistic survey of New Caledonia, the Solomon Islands, and both the Australian- and Dutch-administered portions of New Guinea, in which areas he collected field materials for scores of Austronesian languages. It was this experience perhaps more than any other which committed him to the study of Austronesian linguistics, and more particularly to the study of the processes of linguistic change in Melanesia.

Robert Blust, ed. *Currents in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, 1-4. *Pacific Linguistics*, C-117, 1991.

After completing his fieldwork in the Pacific George spent the year 1956-1957 as an Associate in Malayo-Polynesian Linguistics at the Bishop Museum in Honolulu. In 1958, despite the untoward incident alluded to above, he completed his doctorate under Joseph Greenberg at Columbia University. From 1958 to 1959 he was Assistant Professor of Sociology at the Women's College of the University of North Carolina; from 1959 to 1960 Assistant Professor of Anthropology at Northwestern University; from 1960 to 1963 Assistant Professor of Anthropology, and from 1963 to 1964 Associate Professor of Anthropology at Southern Illinois University. In 1964 he arrived in Hawai'i as a Scholar in Residence at the East-West Center. Before he had completed his term with the East-West Center he was hired as Professor of Linguistics by the University of Hawai'i, where he joined the newly formed Department of Linguistics, headed by Howard McKaughan.

George Grace's professional career encompasses at least four distinguishable roles: those of editor, administrator, scholar and teacher. In 1961, while he was still employed by Southern Illinois University at Carbondale, George became editor of the new journal, *Oceanic Linguistics*. The first four issues of this journal, dedicated to the study of Austronesian, Papuan, and Australian languages, were modest in scope and cost. George carried the editorship of *Oceanic Linguistics* with him to Hawai'i. This transition of employment is reflected in the journal, which in 1964 graduated from stapled to bound format, grew noticeably in size, and began increasingly to attract the contributions of leading scholars concerned with languages of the Pacific area. George has contributed both articles and review articles to *Oceanic Linguistics*, but above all he has remained its editor for the past thirty years (making his editorial tenure probably one of the longest on record). During this time the journal has grown into the primary forum for publications concerned with the Austronesian languages, and one of the major outlets for publications on both the Papuan and the Australian languages.

From 1966 to 1969 George headed the Department of Linguistics at the University of Hawai'i after its first leader, Howard McKaughan, moved into a deanship. I have never spoken to him about the matter, but have always had the impression that being department chair was not his favourite position in the world of linguistics.

George's publications include descriptive studies, most notably dictionaries of two of the languages of New Caledonia, comparative studies of a substantive nature, and works of a more theoretical-philosophical character which cover a range of topics from the nature of language change to the nature of language as an object amenable to scientific study, translation theory and the relationship of language to thought. The interconnectedness of this body of work may initially elude some readers, but I believe that a basic unity of purpose underlies the seeming diversity of interests reflected in it.

George's comparative work has two salient foci. The first of these centres on issues of subgrouping. The second focus concerns the nature of language change, and in particular the issue of how the languages of Melanesia can be divided impressionistically into two groups that he has charmingly compared to birds (drastically altered in the transition from an archosaurian ancestor) and crocodilians (barely changed in the transition from the same ancestor). Any Oceanic linguist who has read his unforgettable discussion of the subject can readily classify, for example, Aneityum as a bird and Fijian as a crocodilian. And, at the same time, anyone who knows George will be aware how far he is willing to push the comparison, given his general distrust of rigid categorisation.

George's dissertation dealt with issues of subgrouping, and he has returned periodically to problems in the genetic classification of the Oceanic languages. His interests in this area have been both substantive, as in his 'Subgrouping of Malayo-Polynesian: a report of tentative findings' (1955)

and *The position of the Polynesian languages within the Austronesian (Malayo-Polynesian) language family* (1959), and theoretical, as in his 'Problems in Oceanic linguistic subgrouping' (1962), 'On the scientific status of genetic classification in linguistics' (1965) or 'Oceanic subgrouping: retrospect and prospect' (1985). Indeed, it is one of the hallmarks of George's work that he rarely discusses issues of substance without at the same time examining the assumptions which underlie their interpretation. A good example is his 'Austronesian lexicostatistical classification: a review article' (1966), a meticulous dissection of the conceptual basis of Isidore Dyen's *A lexicostatistical classification of the Austronesian languages* (1965), and in my view one of the finest review articles ever written in the field of historical linguistics.

It is this tendency to lay bare the philosophical underpinnings of linguistic arguments that provides a bridge between George's earlier work, which is overtly substantive (but often 'philosophical' by predisposition) and his later work in translation and world-view (collectively 'ethnolinguistics'), which is overtly concerned with an examination of the most basic assumptions underlying contemporary theories of language. From my days as a graduate student studying Austronesian linguistics with George I can vividly recall his frequent, only half-facetious references to the 'good', or 'well-behaved' Melanesian languages (= the crocodilians; i.e. those which show relatively straightforward sound changes and respectable cognate densities with other Austronesian languages) and the 'bad', or 'aberrant' Melanesian languages (= the birds; i.e. those which show extreme lexical and phonological divergence which does not seem to be connected in any very straightforward way with differences of separation time). I suspect that it was his desire to come to grips with the reasons for such differences among languages which appear to subgroup together that led George into the philosophical concerns of his later years: What are the units of language structure and of language change? What is the relationship between a language and 'its linguistic description'? What is the relationship between language and thought?

The foregoing are not commonly asked questions in linguistics. Indeed, one of George's most distinctive and engaging characteristics is his incorrigible indifference to fads and fashions. Let me hasten to add that such a statement can easily be misunderstood. George is exceptionally well read, and keeps abreast of the general literature in linguistics and related disciplines far more conscientiously than most of his colleagues. But his research interests are self-driven, rather than inspired and guided by the light out of the East. By asking the questions that intrigued him rather than the questions that were fashionable to ask he has simply charted a course of his own. To a large extent the questions that George has tended to find meaningful in his career have been concerned in one way or another with the integration of synchrony and diachrony. As a linguist who entered the field in the 1950s with strong interests in problems of language change, he has had relatively little to say about syntax, and for this reason his searching examination of the foundations of linguistics in his books *An essay on language* (1981) and *The linguistic construction of reality* (1987) has yet to link up in any decisive way with the main thrust of theoretical work in our field over the past three decades.

It is not easy to summarise in a few words the goals that underlie George's work in linguistic theory, and I am not at all sure that I am qualified at this point to do so. Perhaps most fundamentally, he appears to ask his readers to reconsider the almost universally unquestioned view that linguistics must properly be concerned with language as code (or form) and not at all with language as message (or content). The pivotal term relating to the new perspective he introduces is *content form*. The content form of an utterance is "The way in which the **idea** which it expresses is analyzed (**construed**) for expression – the way it is put into words. This construction in fact creates a model

of a bit of reality (or *as-if* reality)" (*An essay on language*, p.172). From here the train of investigation leads naturally into the role of communicative competence in grammatical description, into problems of translation and translatability, and inevitably into issues of the kind most commonly associated with the name of Benjamin Lee Whorf. To the average working linguist George's discussions of language may appear closer to philosophy of language than to linguistics, while to the average working philosopher of language the reverse may appear true. In short, George Grace is a pioneer who is blazing his own trail in the wilderness between the disciplines.

When I studied with George in the late 1960s it was common for graduate students to be almost overawed by the breadth and depth of his knowledge. In the words of one of these students he was "amazing Grace", and to many of his colleagues George was considered something of a scholar's scholar. In committee work, thesis supervision, or general comments in conversation he has a way of penetrating quickly to the heart of things and expressing solutions in clear and simple terms. Perhaps for this reason as much as any other George has probably been more sought after as a 'general' member of dissertation committees than anyone else in his department (39 times between 1966 and 1990). Not being committed to any particular theoretical doctrine, yet never satisfied with poorly thought-out arguments, he is an ideal sounding-board for any kind of intellectual proposal. At the same time George has chaired a number of dissertation committees, including (in chronological order) those for John Lynch, myself, Kay Ilkanagara, Sheldon Harrison, Frantisek Lichtenberk, Joel Bradshaw, Suzanne Scollon, Amara Prasithrathsint, Anne Pakir, and Hiroshi Sugita. When George presented the keynote address at the symposium on Austronesian linguistics which was held in conjunction with the XV Pacific Science Congress at Dunedin, New Zealand in 1983, it was pointed out that he had been instrumental in training an entire generation of Oceanic linguists. What had been a tiny and arcane discipline in the 1950s had become an arena of lively debate between a number of well-trained, independent scholars in the 1980s, virtually all of whom had been touched directly or indirectly by George's teaching.

I would be remiss to leave the unbalanced impression that George Grace is nothing more than an intellectual giant who has made major and lasting contributions to several branches of linguistics. He is also a former tennis champion (he and his doubles partner Robert McGlone were ranked second in the state of Hawai'i for their age bracket in the 1970s; he repeated this distinction with doubles partner Frank Miller as recently as 1986). Finally, George probably is one of Hollywood's greatest missed opportunities for a challenger to Woody Allen. Anyone who has ever heard a Gracian public presentation, with the inevitable opening barrage of disclaimers, apologies, advance qualifications, etc., will have experienced that impossible mixture of emotions I have felt – not knowing whether to rush up to offer him one's assurances, or to laugh out loud. Most people who work with him on a daily basis have a very hard time keeping a straight face in talking to him (I have always made a special effort). In the subtlest and most inimitable way George gives one the distinct impression that he believes the world is out to get him, and that his only defence is humour. If this is so, he defends himself extraordinarily well.



THE PUBLICATIONS OF GEORGE W. GRACE

COMPILED BY ALBERT J. SCHÜTZ

1955

Subgrouping of Malayo-Polynesian: a report of tentative findings. *American Anthropologist* 57:337-339.

1958

Review of H.R. Klieneberger *Bibliography of Oceanic linguistics*, London, 1957. *American Anthropologist* 60:1243.

1959

*The position of the Polynesian languages within the Austronesian (Malayo-Polynesian) language family*. Indiana University Publications in Anthropology and Linguistics, Memoir 16; supplement to *International Journal of American Linguistics* 25; also, Bernice P. Bishop Museum Special Publication 46. Baltimore: Waverly Press. (His PhD dissertation, Columbia University, 1958.)

1960

(with A.L. Kroeber) *The Sparkman grammar of Luiseño*. University of California Publications in Linguistics 16. Berkeley: University of California Press.

1961

Abstract: Lexicostatistics and Austronesian subgrouping. In: Abstracts of symposium papers, Tenth Pacific Science Congress, Honolulu, 96-97.

Austronesian linguistics and culture history. *American Anthropologist* 63:359-368. (Reprinted in Harding and Wallace, eds, 1970 - see below.)

Lexicostatistical comparison of six Eastern Austronesian languages. *Anthropological Linguistics* 3/9:1-22.

Review of K.J. Hollyman *A checklist of Oceanic languages*, Auckland, 1960. *Journal of the Polynesian Society* 70:378-379.

1961-

(editor) *Oceanic Linguistics*. (GWG has edited *Oceanic Linguistics* from its foundation (following the Tenth Pacific Science Congress in 1961) to date.

1962

Comment on Arthur Capell, Oceanic linguistics today. *Current Anthropology* 3/4:408-410.

Problems in Oceanic linguistic subgrouping. *Transactions of the New York Academy of Sciences*, series 2, 24:684-689.

1963

Abstracts and translations from articles on Oceanic linguistics. *International Journal of American Linguistics* 29:153-157.

Review of Arthur Capell *A linguistic survey of the South-Western Pacific*, Nouméa, 1962. *American Anthropologist* 65:481-482.

1964

Abstracts and translations from articles on Oceanic linguistics.. *International Journal of American Linguistics* 30:175-178.

The linguistic evidence. In: Kwang-chih Chang, George W. Grace and William G. Solheim II, Movement of the Malayo-Polynesians: 1500 B.C. to A.D. 500. *Current Anthropology* 5:361-368, 403-404.

1965

On the scientific status of genetic classification in linguistics. *Oceanic Linguistics* 4:1-14.

Review of Dell Hymes, ed. *Language in culture and society: a reader in linguistics and anthropology*, New York, 1964. *Social Forces* 44:289.

Review of H.L. Shorto, ed. *Linguistic comparison in South East Asia and the Pacific*, London, 1963. *American Anthropologist* 67:1019-1020.

1966

Austronesian lexicostatistical classification: a review article. *Oceanic Linguistics* 5/1:13-31. (Comments by Dyen, and Hymes, in succeeding pages.)

1967

Effect of heterogeneity in the lexicostatistical test list: the case of Rotuman. In: Genevieve A. Highland et al., eds *Polynesian culture history: essays in honor of Kenneth P. Emory*, 289-302. Bernice P. Bishop Museum Special Publication 56. Honolulu: Bishop Museum Press.

1968

Classification of the languages of the Pacific. In: Andrew P. Vayda, ed. *Peoples and cultures of the Pacific: an anthropological reader*, 63-79. Garden City, New York: The Natural History Press.

Review of Jerry Allen and Conrad Hurd *Languages of the Bougainville District*, Port Moresby, 1965. *Lingua* 19/4:437-441.



Review of G.B. Milner and Eugénie J.A. Henderson, eds *Indo-Pacific linguistic studies* (*Lingua* 14 & 15), Amsterdam, 1965. *Language* 44:168-172.

1969

Notes on the philosophical background of current linguistic controversy. *Working Papers in Linguistics, University of Hawaii* 1/1:1-41.

A Proto-Oceanic finder list. *Working Papers in Linguistics, University of Hawaii* 1/2:39-84.

Speaking of linguistic change. *Working Papers in Linguistics, University of Hawaii* 1/3:101-116.

1970

Austronesian linguistics and culture history. In: Thomas G. Harding and Ben J. Wallace, eds *Cultures of the Pacific*, 20-28. New York: The Free Press. (Reprinted from *American Anthropologist*, 1961 - see above.)

Languages of Oceania. *Working Papers in Linguistics, University of Hawaii* 2/3:1-24.

Notes on the phonological history of the Austronesian languages of the Sarmi Coast. *Working Papers in Linguistics, University of Hawaii* 2/9:55-104. (Reprinted in *Oceanic Linguistics*, 1971 - see below.)

Oceanic linguistics tomorrow. In: S.A. Wurm and D.C. Laycock, eds *Pacific linguistic studies in honour of Arthur Capell*, 31-65. *Pacific Linguistics*, C-13.

On canonical shapes in the language of Koumac. *Working Papers in Linguistics, University of Hawaii* 2/5:205-222. (Reprinted in Thomas and Bernot, eds, 1972 - see below.)

1971

Languages of the New Hebrides and Solomon Islands. In: Thomas A. Sebeok, ed. *Current trends in linguistics*, vol.8: *Linguistics in Oceania*, 341-358. The Hague: Mouton. (GWG was an Associate Editor of volume 8 of *Current trends in linguistics*.)

Notes on the phonological history of the Austronesian languages of the Sarmi Coast. *Oceanic Linguistics* 10/1:11-37. (Reprinted from *Working Papers in Linguistics, University of Hawaii*, 1970 - see above.)

1972

On canonical shapes in the language of Koumac. In: Jacqueline M.C. Thomas and Lucien Bernot, eds *Langues et techniques, nature et société*, vol.1: *Approche linguistique*, 135-140. Paris: Klincksieck.

1973

Foreword, to *Oceanic Linguistics* 12/1-2:v.

(et al., eds) *Papers of the First International Conference on Comparative Austronesian Linguistics, 1974 - Oceanic*. *Oceanic Linguistics* 12/1-2. (Appeared 1975.)

Research on the position of the New Caledonian languages: a progress report. *Working Papers in Linguistics, University of Hawaii* 5/7:49-62.

Review of Earl J. Hughes *Nunggubuyu-English dictionary*, Sydney, 1971. *American Anthropologist* 75/2:502-503.

1974

(et al., eds) *Papers of the First International Conference on Comparative Austronesian Linguistics, 1974 - Proto-Austronesian and Western Austronesian*. *Oceanic Linguistics* 13/1-2. (Appeared 1976.)

1975

*Canala dictionary (New Caledonia)*. *Pacific Linguistics*, C-2.

Linguistic diversity in the Pacific: on the sources of diversity. *Working Papers in Linguistics, University of Hawaii* 7/3:1-7.

1976

*Grand Couli dictionary (New Caledonia)*. *Pacific Linguistics*, C-12.

History of research in Austronesian languages of the New Guinea area: general. In S.A. Wurm, ed. *New Guinea area languages and language study*, vol.2: *Austronesian languages*, 55-71. *Pacific Linguistics*, C-39. (Reprinted in Lynch, ed., 1981 - see below.)

Review article: Andrew Pawley, On the internal relationships of the Eastern Oceanic languages, in: R.C. Green and M. Kelly, eds *Studies in Oceanic culture history* 3, Honolulu, 1972. *Journal of the Polynesian Society* 85:103-112.

1978

Frantisek Lichtenberk's sketch of Houailou grammar. *Working Papers in Linguistics, University of Hawaii* 10/2:73.

Introduction [to volume of Papers from the Oceanic Comparative Linguistics Project]. *Working Papers in Linguistics, University of Hawaii* 10/1:1-7.

1979

The dilemma of the disciplines. *Working Papers in Linguistics, University of Hawaii* 11/1:1-10.

On "linguistic change" and related notions. *Working Papers in Linguistics, University of Hawaii* 11/3:27-34.

On "meaning" and related notions. *Working Papers in Linguistics, University of Hawaii* 11/3:35-46.

1980

Comment on Robert Blust, Early Austronesian social organization: the evidence of language. *Current Anthropology* 21/2:235.

1981

*An essay on language*. Columbia, South Carolina: Hornbeam Press.

History of research in Austronesian languages of the New Guinea area: general. In John Lynch, ed. *Readings in the comparative linguistics of Melanesia*, 67-77. Studying Pacific Languages Series, no.4. Waigani: Language Department, University of Papua New Guinea. (Reprinted from *Pacific Linguistics*, C-39 - see above.)

Indirect inheritance and the aberrant Melanesian languages. In: Jim Hollyman and Andrew Pawley, eds *Studies in Pacific languages & cultures in honour of Bruce Biggs*, 255-268. Auckland: Linguistic Society of New Zealand.

Review article: S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, Canberra, 1978. *Journal of the Polynesian Society* 90/3:426-430.

1982

Thoughts on translation and meaning: a progress report. *Working Papers in Linguistics, University of Hawaii* 14/3:1-23.

1983

Review of S.A. Wurm and Shirô Hattori, eds *Language atlas of the Pacific area*, Canberra, 1981, 1983. *Asian Perspectives* 25/2:139-141 (appeared in 1988).

1985

Oceanic subgrouping: retrospect and prospect. In Andrew Pawley and Lois Carrington, eds *Austronesian linguistics at the 15th Pacific Science Congress*, 1-18. *Pacific Linguistics*, C-88.

On the explanation of sound changes: some Polynesian cases. In Veneeta Z. Acson and Richard L. Leed, eds *For Gordon H. Fairbanks*, 56-63. *Oceanic Linguistics* Special Publication 20. Honolulu: University of Hawaii Press.

Review of Eric de Grolier, ed. *Glossogenetics: the origin and evolution of language*, Chur/New York, 1983. *Anthropological Linguistics* 27: 112-117.

1986

Further thoughts on Oceanic subgrouping. In: 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.

Hypotheses about the phonological history of the language of Canala, New Caledonia. In: Chris Come and Andrew Pawley, eds *Le coq et le cagou: essays on French & Pacific languages in honour of Jim Hollyman*, 55-67. *Te Reo* 29. Auckland: Linguistic Society of New Zealand.

1987

Genetic classification revisited. In: Donald C. Laycock and Werner Winter, eds *A world of language: papers presented to Professor S.A. Wurm on his 65th birthday*, 221-226. *Pacific Linguistics*, C-100.

*The linguistic construction of reality*. London: Croom Helm.

1988

Do languages change at a constant rate: an alternative view of basic vocabulary replacement. In: Roger L. Hadlich and J.D. Ellsworth, eds *East meets West: homage to Edgar C. Knowlton, Jr*, 72-79. Honolulu: Department of European Languages and Literature, College of Languages, Linguistics, and Literature, University of Hawaii.

1989

The meaning of meaning: review article of Mark Johnson *The body in the mind: the bodily basis of meaning, imagination, and reason*, Chicago, 1987. *Semiotica* 73:351-362.

1990

The "aberrant" (vs. "exemplary") Melanesian languages. In: Philip Baldi, ed. *Linguistic change and reconstruction methodology*, 155-173. Berlin: Mouton de Gruyter.

"Consonant grade" in Oceanic languages. In: J.H.C.S. Davidson, ed. *Pacific island languages: essays in honour of G.B. Milner*, 41-49. London: School of Oriental and African Studies, University of London/Honolulu: University of Hawaii Press.

in press

The university teaching of linguistics in Hawaii. In Thomas A. Sebeok and John H. Hammer, eds *University teaching of linguistics in the United States*. Mouton.

Languages of Oceania. *The Linguistic Reporter*.

submitted

Syntactic change in progress? [in volume in memory of Fang-kuei Li].

## ON THE CATEGORY DISTRIBUTIVE

BYRON W. BENDER

### 1. INTRODUCTION

The term **distributive** has been used on occasion in the description of languages that present a formally distinct category to which it seems to apply. Latin grammars, for example, present the special distributive forms of numerals which serve to answer the question 'How many of each?' or 'How many at a time?' Thus *singula singulis* is translated "one apiece (one each to each one)" (Greenough et al. 1903:61). The term has found sporadic use in some Austronesian languages for other grammatical phenomena beyond the numerals, especially in certain Philippine languages. For example, Blake (1925:58) says that some fully reduplicated roots in Tagalog combined with *mag--an* may take the infix *in* in the first syllable of the root "with a distributive meaning," and gives the one example:

*maghinanaphanapán* 'to seek each his own'.

For the Aklanon language, de la Cruz and Zorc (1968:106) state that a particular *pang-* prefix and its accompanying morphophonemic changes give verbs a distributive quality:

Instead of a simple or general statement of the action, the action is distributed through time, or among several people or objects, or is made into a process. The changes that take place usually involve local idiom, and are not always predictable.

To cite three of their examples, *hueog* 'drop' becomes *-panghueog* 'drop plenty of things', *pasyar* 'promenade' becomes *-pamasyar* 'walk around', and *asawa* 'spouse' becomes *-pangasawa* 'court and marry'.

Beyond brief mention or short sections in grammars such as these, I am not aware that **distributive** as a label for a grammatical category has received any extensive or general discussion in the literature. In this paper I will attempt an initial step in this direction, based on some facts from Marshallese, a language in which distributive phenomena figure prominently.<sup>1</sup>

---

<sup>1</sup>I am indebted to two of the people who have taught me much of what I know about the language and culture of the Marshall Islands, Alfred Capelle and Tony DeBrum, for assisting me at various turns with this paper. My thanks also to Louise Pagotto and Ken Rehg for their helpful comments on an earlier version.

Robert Blust, ed., *Currents in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, 11-26.  
*Pacific Linguistics*, C-117, 1991.

© Byron W. Bender

## 1.1 PREVIOUS WORK

The language lessons and grammatical notes of Bender (1969), through examples and informal discussion, give something of the range and diversity of both form and meaning associated with the distributive in Marshallese. Pagotto (1987), using an expanded database that includes the example sentences from Abo et al. (1976), formalises three derivational rules (DR), all of which derive verbs of one sort or another. Her DR-14 derives stative distributive verbs from common nouns:

d-f- $\emptyset^2$	<i>bukwōn</i>	'chunk'
	<i>bbukwōnkwōn</i>	'be chunky' (pp.586-589)

DR-18 adds the semantic feature 'distributed' to a variety of intransitive verb subclasses:

d-f- $\emptyset$	<i>kōto</i>	'wind'
	<i>kkōtoto</i>	'always be windy' (pp.598-602)

DR-19 adds the same distributed semantic feature to transitive verbs:

$\emptyset$ -i-e	<i>deñōt</i>	'spank someone'
	<i>deñdeñōte</i>	'spank all over someone' (pp.603-605)

The first two of these examples follow the same pattern, one which employs initial-consonant doubling (d) and final-syllable reduplication (f); the pattern of the last example employs initial syllable reduplication (i) and -e suffixation (e).<sup>3</sup> These abbreviations are used in the labels I have given the patterns above, with zero being used as a place-holder for a non-occurring element. (Initial- and final-syllable reduplication do not co-occur in the same pattern, and can thus both be accorded the middle slot in the labels given the patterns.)

Each of the three rules as presented by Pagotto has several morphological patterns associated with it – from three to five – with no apparent conditioning.

In addition to the **d-f- $\emptyset$**  pattern illustrated above, DR-14 may have either  **$\emptyset$ -f-e** or  **$\emptyset$ - $\emptyset$ -e**:

$\emptyset$ -f-e	<i>jɔɔ!</i>	'salt'
	<i>jɔɔ!ɔ!e</i>	'be salty'
$\emptyset$ - $\emptyset$ -e	<i>dān</i>	'water'
	<i>dāne</i>	'be juicy'

There are four alternatives to the **d-f- $\emptyset$**  pattern illustrated above for DR-18:

$\emptyset$ -f-e	<i>eɔɔt</i>	'be striped'
	<i>eɔɔtɔte</i>	'be striped all over'
d- $\emptyset$ - $\emptyset$	<i>jālele</i>	'eat sauce'
	<i>jjālele</i>	'always eat sauce'
$\emptyset$ -f- $\emptyset$	<i>badik</i>	'bow down'
	<i>badikdik</i>	'humble oneself'

<sup>2</sup>The key for these abbreviations is given in the first paragraph following DR-19 below.

<sup>3</sup>With a few exceptions, the -e suffix has the variant -i when the preceding vowel is high. Some of the exceptions are those that manifest the alternation to be found in *iɔwe* 'have many houses', *ike* 'be teeming with fish', and *arnije* 'be inhabited' (cf. respectively *epn* 'house', *ek* 'fish' and *arnej* 'people'). Other exceptions such as *jidjide* 'lots of crickets' (cf. *jidjid* 'cricket') do not manifest the alternation, but may nevertheless relate to the same historical phenomena, for a discussion of which see Bender (1984:449-451).



$\emptyset$ - $\emptyset$ -e	<i>bok</i>	'blister'
	<i>boke</i>	'have blisters all over'

There are three alternatives to the  $\emptyset$ -i-e pattern illustrated above for DR-19:

d-i-e	<i>rup</i>	'break something'
	<i>rruprupe</i>	'break something carelessly'
d-i- $\emptyset$	<i>jibwe</i>	'hold someone/something'
	<i>jjibjibwe</i>	'fondle someone'
$\emptyset$ -i- $\emptyset$	<i>jibwe</i>	'hold someone/something'
	<i>jibjibwe</i>	'fondle someone'

Occasionally a given verb may alternate among patterns, resulting in variant forms, as in the last two examples.

Following is a summary of the alternative patterns Pagotto builds into her rules. Note that there is considerable overlap for the morphological patterns of DR-14 and DR-18, but complete complementarity between them, on the one hand, and DR-19 on the other. Whereas the former may involve final-syllable reduplication (f), the latter never does, and requires instead initial-syllable reduplication (i).

DR-14	DR-18	DR-19
Derive:		
N $\rightarrow$ V <sub>st</sub> distr	V <sub>intr</sub> $\rightarrow$ V <sub>intr</sub> distr	V <sub>tr</sub> $\rightarrow$ V <sub>tr</sub> distr
Morphological patterns:		
d-f- $\emptyset$	d-f- $\emptyset$	d-i- $\emptyset$
$\emptyset$ - $\emptyset$ -e	$\emptyset$ - $\emptyset$ -e	--
$\emptyset$ -f-e	$\emptyset$ -f-e	$\emptyset$ -i-e
--	$\emptyset$ -f- $\emptyset$	$\emptyset$ -i- $\emptyset$
--	d- $\emptyset$ - $\emptyset$	--
--	--	d-i-e

Concerning the range of semantic distinctions associated with the distributed feature, Pagotto does not go into detail beyond noting that:

the semantic relationships between the source and derived distributive are not always predictable. While many distributive verbs have the sense of 'spread out, all over, always', some take on a continuative meaning..., while others have an idiosyncratic relationship....(p.601)

She also notes that while

the derived transitive verb is often interpreted as 'do something carelessly, all over the place'.... [it is] not limited to this interpretation, e.g. *ttūmtūmi* 'keep pulling s.'..... (p.605)

## 1.2 THE PRESENT WORK

This short paper does not attempt to find conditioning factors for the various morphological patterns or predictability for the seemingly idiosyncratic semantics. At most it is my hope to add a bit more perspective, and possibly bring some unity to what seems to be a most disparate category, if indeed it is a category at all. One part of Pagotto's treatment I do hope to be able to flesh out is the relative frequency of each morphological pattern.

This I attempt to do by looking at the phenomena as exhaustively as possible. My starting point is Abo et al. (1976). Early in the compilation of this dictionary we became aware that for many of the basic entries there were derived forms with a distributive sense, and we proceeded to check each entry systematically for this potential. In the end, special derived distributive forms were recorded for 860 of the almost 5000 entries, roughly one out of six. Example sentences for a number of these derived forms were also collected then, and remaining gaps have since been filled for the present study.<sup>4</sup>

I have proceeded to give each distributive a three-way classification according to the major lexical category that serves as its base, and how it differs both formally and semantically from that base. Bases that were loanwords were also noted. This classification was then sorted in various ways to bring out relevant intersections, which are reported in what follows.

## 2. FORM

The following are approximate numbers of the incidence of each morphological pattern found for each of Pagotto's rules:

DR-14	V <sub>st</sub> distr	DR-18	V <sub>intr</sub> → V <sub>intr</sub> distr	DR-19	V <sub>tr</sub> → V <sub>tr</sub> distr
Derive:					
N →		V <sub>intr</sub> →		V <sub>tr</sub> →	
Morphological patterns and frequencies: <sup>5</sup>					
d-f-∅:	150	d-f-∅:	370	d-i-∅:	10
∅-∅-e:	60	∅-∅-e:	35	---	
∅-f-e:	10	∅-f-e:	2	∅-i-e:	5
∅-f-∅:	15	∅-i-∅:	45	∅-i-∅:	5
d-∅-∅:	5	d-∅-∅:	30	---	
---		---		d-i-e:	10

It will be noted that identical sets of patterns are now shown for the first two rules.<sup>6</sup> That is, whereas Pagotto did not show either ∅-f-∅ or d-∅-∅ as possible morphological patterns for DR-14,

<sup>4</sup>I am indebted to dictionary co-author Alfred Capelle for supplying some 300 additional example sentences early in 1990, so that each of the 860 attested distributives is now exemplified by at least one sentence.

<sup>5</sup>The frequencies given here should not be expected to total 860. As will be seen in the discussion to follow, other patterns were found, and for the inherent distributives (also discussed below), there are no derivational rules.

<sup>6</sup>Since the morphological patterning of DR-14 and DR-18 proves to be identical, it may seem tempting to conflate them into one rule. However, to do so would violate Aronoff's Unitary Base Hypothesis (1976:47-48), since DR-14 involves nouns, and DR-18 verbs.



a full examination of the data used for this paper shows 15 examples of the former and five of the latter. Here is an example of each:

$\emptyset$ -f- $\emptyset$	<i>di</i>	'bone'
	<i>didi</i>	'lots of little bones'
d- $\emptyset$ - $\emptyset$	<i>naan</i>	'word'
	<i>nnaan</i>	'news'

Before leaving this discussion of form, it should be noted that for both non-transitive rules, the pattern with the highest frequency by far is **d-f- $\emptyset$** , which involves both initial-consonant doubling and final-syllable reduplication. In fact, some instances of single **d- $\emptyset$ - $\emptyset$**  and  **$\emptyset$ -f- $\emptyset$**  can be seen as serving to accomplish this majority pattern, when the base already contains one of the two elements. That is, for a base that has an inherent initial double consonant (d) such as *mpaan* 'man, male', final-syllable reduplication (f) alone serves to accomplish the full **d-f- $\emptyset$**  pattern: *mpaanan* 'manly; lots of men'. Similarly, (d) alone serves the same end for a base with inherent final-syllable reduplication (f): *likaeb* 'cone shell', *likaebeb* 'lots of cone shells'.

Not only is the **d-f- $\emptyset$**  pattern predominant for the first two rules, but an examination of distributive formation for loanwords shows, as might be expected, that it is also the currently productive pattern. Of some 90 loans observed in the data, all follow this pattern, with only the following exceptions:

<i>anien</i>	'onion'
<i>aniene</i>	'lots of onions'
<i>kūrij</i>	'fatty meat, blubber, lard'
<i>kūriji</i>	'be excessively greasy, of food'
<i>eļbōn</i>	'elephant'
<i>eļbōne</i>	'be teeming with elephants'
<i>jiip</i>	'jeep'
<i>jiipipi</i>	'be obsessed with jeeps' (d-f-e)

Occasionally, as in this last example, the two major patterns are combined: **d-f- $\emptyset$**  and  **$\emptyset$ - $\emptyset$ -e** to yield **d-f-e**. But in general, the pattern of second highest frequency,  **$\emptyset$ - $\emptyset$ -e**, (with the few seeming exceptions listed here) seems to be no longer productive.<sup>7</sup>

Other bits of evidence attesting to the strength and productivity of the **d-f- $\emptyset$**  pattern are the instances of its being used with DR-19 to derive transitive distributives, such as those given below. This number, still relatively small but no doubt increasing, is working to destroy the complementarity between initial- and final-syllable reduplication Pagotto found. (There is little doubt that initial-syllable reduplication was once the prevailing pattern, as presented by Pagotto.)

d-f- $\emptyset$	<i>kūtimi</i>	'wrap something'
	<i>kkūtimtimi</i>	'wrap something all the time'

<sup>7</sup>One additional partial exception worth mentioning concerns the doublet formed from the root *mpakpōk*. One contemporary lexeme on this root refers to the arrowroot plant, and its distributive manifests the same alternation referred to in Bender 1984:449-451: *mpakpūke* 'be thick with arrowroot plants'. Another, referring to the starch product from the plant, and possibly of more recent origin, simply copies the quality of the vowel of the preceding syllable in the suffix of its distributive: *mpakpōke* 'be starchy'. An example sentence for the latter is given under *sufficiency* below.

d-f-∅	<i>tōnōle</i>	'poke someone/something'
	<i>ttōnōlñōle</i>	'poke someone/something excessively'
d-f-∅	<i>kawūnouk(i)</i>	'dye something'
	<i>kkawūnonouki</i>	'spread dye out on something'

Several additional examples are given under the discussion of meaning that follows.

Before leaving the discussion of form, it should be mentioned that there are occasional irregularities in distributive formation, such as the following, which depart sporadically from the patterns that have been identified:

<i>wōt</i>	'rain'
<i>wūttuot</i>	'lots of rain'
<i>owe</i>	'to whistle'
<i>ajjowewe</i>	'to keep on whistling'

### 3. NUMERALS

The formation of distributives from numerals conforms only in part to the formal and semantic patterns of other distributives. Pagotto (1987) does not include this in her study. She does, however, identify numerals and other quantifiers as a subclass of noun (p.151). If DR-14 were to operate on the numerals using the predominant **d-f-∅** pattern, *\*jjilulu* would be produced from *jilu* 'three', for example. Although this form does not occur as such, what we find is the same form preceded by the causative prefix: *kajjilulu* 'keep taking three each'. A parallel causative derivation using existing bases (including distributives) is formalised by Pagotto (p.607) as DR-20. Distributives are regularly formed from numerals in this manner, which can be symbolised as **ka-d-f-∅**.<sup>8</sup> This will be referred to as the continuative distributive for numerals.

Two other distributives are regularly formed from the numerals. One is similar to that just described, but without final-syllable reduplication (and for the numerals from one to five, optional doubling of the second consonant as well). Thus, from base *jilu* the form is *kajji(l)lu* 'take three each', with the meaning contrast being the absence here of an inherent continuative aspect. The form pattern for this distributive numeral can be symbolised as **ka-d-∅-∅**.

The third regular distributive can be viewed as derived from the second by a transitivity rule (Pagotto's DR-31 (p.640)). Thus for 'three' it would yield *kajjiluuki* 'take three each from someone'.

### 4. INHERENT DISTRIBUTIVES

Not all distributives are derived, it would seem. There are some 20 instances in the database that are distributive in both form and meaning, but for which no source (an underived base) can be found.

<sup>8</sup>The one exception is *juon* 'one', which first loses what is probably a fossilised *-n* suffix. This original root *\*juo* is then further modified according to a regular alternation that accompanies final vowel apocope to yield the high-mid vowel that occurs in the reduplicated syllable of *kajjojo* [kajjéwjéw] 'distribute into ones'.

<i>ppedikdik</i>	'be slow, tardy'	d-f-∅
<i>Jab ppedikdik bwe eboñ.</i> [not ~ because 3SG-be.night] <sup>9</sup> 'Don't be slow because it's almost night'.		
<i>bbałokłok</i>	'to bulge'	d-f-∅
<i>Ebbałokłok eoon tebōłeɣ.</i> [3SG- ~ on-of table that] 'The top of that table is bulgy'.		
<i>ɣɣōłeiññ</i>	'to bob, rock, roll, sway'	d-f-∅
<i>ɣɣōłeiññ in wa eo ekōɣōłañłōñe.</i> [~ of canoe the 3SG-CAUS-nausea-TR] 'The swaying of the ship made her want to vomit'.		

## 5. MEANING

Let us turn now to the meaning distinctions one finds in distributives as compared with their underived bases. In the sections that follow I give the rubrics I have found useful in my attempt at semantic classification, together with a rough indication of the size of each group. Several examples are given for each group, some chosen for their representativeness, others because their example sentences reveal relations to other groups.

What we find in the semantics of the distributive is a constellation of related meanings. They are not all clearly distinct from each other; often they seem to shade into each other, especially their closest neighbours. Some of them I have found it necessary to reclassify more than once. In spite of this seeming indeterminacy and lack of discreteness, the task proves worthwhile, as will be seen.

### PERVASIVE (250)

Near the heart of the constellation is the group I have labelled as being pervasive. The quality of the base spreads to fill a space (the bounds of which may or may not be specified) without dilution.

<i>kiaj</i>	'gasoline'	<i>kkiaajaj</i>	DR-14	d-f-∅
<i>Ekkiaajaj ioon pein bałun eo.</i> [3SG- ~ on-of wing-3SG airplane the] 'There was gas all over the plane's wings'.				
<i>bwiin-jikka</i>	'cigarette odour'	<i>bwiin-jjikkaka</i>	DR-18	d-f-∅
<i>Ebwiin-jjikkaka lowaan ruuɣ in.</i> [3SG- ~ inside-of room this] 'I smell cigarette odour all over this room'.				
<i>tōñal</i>	'to be sweet'	<i>ttōñalñal</i>	DR-18	d-f-∅
<i>Enno an ttōñalñal.</i> [3SG-taste.good thing-3SG ~] 'It's pervasive sweetness is delicious'.				
<i>wōñōt</i>	'be newly constructed'	<i>owōñōtnōt</i>	DR-18	d-f-∅
<i>Ej baj owōñōtnōt wa eo waan.</i> [3SG-PROG just ~ canoe the vehicle-3SG] 'His canoe certainly looked new'.				

The pervasiveness of the last example may not be apparent. But comparison with a parallel sentence with non-distributive verb brings out the distinction that there are a number of things about this canoe that give away its newness. The newness pervades the entire canoe.

<sup>9</sup>I follow the glossing conventions of Pagotto (1987), and also let the tilde stand for the distributive form being exemplified. Among the less transparent grammatical labels used are CMPL completive, IRRL irrealis and PTTV optative.

## SUFFICIENCY (10)

Some distributives seem to go beyond pervasiveness to assert **sufficiency** for a certain purpose.

<i>teaak</i>	'provisions for journey'	<i>teaake</i>	DR-14	Ø-Ø-e
<i>Eteaake wa in.</i> [3SG- ~ canoe this] 'We have plenty of food aboard'.				
<i>walni</i>	'copra, ripe coconut	<i>owainini</i>	DR-14	d-f-Ø
<i>Eowainini ke?</i> [3SG- ~ Q] 'Does it have enough/much grated coconut on it?'				
<i>jiit</i>	'sheet'	<i>jiitit</i>	DR-14	d-f-Ø
<i>Kajjiititi būtoñ ge.</i> [CAUS- ~ mattress that] 'Put enough sheets on the mattress'.				
<i>jikraip</i>	'scribe'	<i>jjikraipip</i>	DR-14	d-f-Ø
<i>Ejjikraipip kweløk eo.</i> [3SG- ~ meeting the] 'The meeting had enough scribes attending'.				
<i>makmōk</i>	'starch'	<i>makmōke</i>	DR-14	Ø-Ø-e
<i>Kōmakmōke iioøk ge arro.</i> [CAUS- ~ mixture that thing-1PL-in-DUAL] 'Put enough starch in our concoction'.				

## FACILITY (30)

The **sufficiency** of the preceding section is all nominally based. Its verbally based complement seems to be focus on ease or susceptibility.

<i>jɔ</i>	'ignite, run, of engines'	<i>jjɔjɔ</i>	DR-18	d-f-Ø
<i>Ejjɔjɔ injin ep.</i> [3SG- ~ engine that] 'That engine is easy to start'.				
<i>tok</i>	'ignite, catch fire'	<i>ttoktok</i>	DR-18	d-f-Ø
<i>Ettoktok kiaaj.</i> [3SG- ~ gasoline] 'Gasoline is flammable'.				
<i>roñ</i>	'to hear'	<i>rroñoñ</i>	DR-18	d-f-Ø
<i>Erroñroñ leep.</i> [3SG- ~ guy-that] 'He has good hearing'.				
<i>meñ</i>	'smell or taste sour'	<i>mmeñmeñ</i>	DR-18	d-f-Ø
<i>Emmeñmeñ jokkwōp in mā.</i> [3SG- ~ soup of breadfruit] 'Breadfruit soup sours quickly'.				
<i>uwōta</i>	'be in danger, afraid'	<i>uwōtata</i>	DR-18	d-f-Ø
<i>Euwōtata leep.</i> [3SG- ~ guy-that] 'That man scares easily'.				
<i>būroṃōj</i>	'be sorry, sad'	<i>bbūroṃōjṃōj</i>	DR-18	d-f-Ø
<i>Ebbūroṃōjṃōj leep</i> [3SG- ~ guy-that] 'He has a tendency toward melancholy'.				
<i>dekā</i>	'yaws'	<i>ddekākā</i>	DR-18	d-f-Ø
<i>Eddekākā ajiri raṃ nejin.</i> [3SG- ~ child those child-3SG] 'Her children are very susceptible to yaws'.				
<i>Ejjeløk wōt ddekākā in laḍik ep.</i> [3SG-not.exist only ~ of boy that] 'That boy is really covered with yaws'.				

## SURFEIT (20)

In some cases, sufficiency and facility are exceeded.

- ijj* 'yeast' *ijjij, ij* DR-14, 18  $\emptyset$ -f- $\emptyset$ ,  $\emptyset$ - $\emptyset$ -e  
*Eijij an iiök pilawā*. [3SG- ~ thing-3SG mix bread] 'The bread she makes smells too much of yeast'.  
*Eijij iiök eo an*. [3SG- ~ mixture the thing-3SG] 'His batter came out having too much yeast'.
- jāde* 'be in sight' *jjādede* DR-18 d-f- $\emptyset$   
*Kwōn jab kōjjādede bwe kwe leddik*. [2SG-PTTV not CAUS- ~ because you girl] 'Don't show yourself in public too much for you're a girl'.
- luwajet* 'dilapidated' *luwajete* DR-18  $\emptyset$ - $\emptyset$ -e  
*Ejab to an pād kōn an luwajete*. [3SG-not long thing-3SG stay because.of thing-3SG ~] 'It didn't stay around very long because it had been in the water too long'.
- llegṇaj* 'jerk' *llegṇajṇaj* DR-18  $\emptyset$ -f- $\emptyset$   
*Etūṇ toon aik eo kōn an llegṇajṇaj*. [3SG-break line-of tow the because.of thing,3SG ~] 'The tow line broke because it was jerked too often'.

## OBSESSIVE (10)

In some cases, all nominally based, the surfeit is not substantive but mental.

- bwidej* 'ground, soil' *bbwidejdej* DR-14 d-f- $\emptyset$   
*Ebbwidejdej leeq*. [3SG- ~ guy-that] 'He's real estate minded'.
- jiip* 'jeep' *jjiiipip* DR-14 d-f- $\emptyset$   
*Ejjiiipipi an kōnono*. [3SG- ~ thing-3SG speak] 'He's always talking about jeeps'.
- juuj* 'shoes' *jjuujuj* DR-14 d-f- $\emptyset$   
*Elukkuun jjuujuj leeq*. [3SG-really ~ guy-that] 'He's preoccupied with shoes'.
- tōṇato* 'tomato' *ttōṇatoto* DR-14 d-f- $\emptyset$   
*Ettōṇatoto aṇ leṇṇak*. [3SG- ~ thing-2SG think] 'You're thinking too much of tomatoes'.

This concludes our tour from the pervasive core of the constellation out to its most excessive reaches. Now let's go back to the core, and begin our exploration in the opposite direction, first along a time dimension.

CONTINUATIVE (150)<sup>10</sup>

Strictly speaking, these are discrete events, yet through hyperbole their unceasing nature is asserted. Note that they may be the products of any of the three derivational rules.

- baid* 'smoking pipe, cigarette' *bbaidid* DR-14 d-f- $\emptyset$   
*Emake bbaidid lōṇe*. [3SG-be.alone ~ guy-that] 'That guy is a chain smoker'.
- baj* 'bus, ride a bus' *bbajbaj* DR-18 d-f- $\emptyset$   
*Kwomake baj bbajbaj*. [2SG-be.alone just ~] 'You're on the bus all the time'.

<sup>10</sup>The term *durative* could probably also be used.

- jepdak* 'crushed in' *jjepdakdak* DR-18 d-f- $\emptyset$   
*Ejjepdakdak neen wa eq waan.* [3SG- ~ foot-of vehicle that vehicle-3SG] 'Her/his vehicle always has flat tires'.  
*Ejjepdakdak kuwat ko.* [3SG- ~ can the-PL] 'The cans are all smashed'.<sup>11</sup>
- daṇwīj* 'lick someone/something' *daṇḍiṇwīj, ddaṇḍeṇwīj*<sup>12</sup> DR-19 (d)-i- $\emptyset$   
*Ear bajjek daṇḍiṇwīj ḷḷe eo.* [3SG-CMPL just ~ candy the] 'He just kept licking the lollipop'.
- kuul* 'squeeze someone/something' *kkukuul, kukuul* DR-19 (d)-i- $\emptyset$   
*Kwōn jab kkukuul bao ɸe.* [2SG-PTTV not ~ bird that] 'Don't keep squeezing that bird'.

The dimension here is time: discrete events are asserted to be as closely packed as possible. We turn now to the analogue in space.

## MULTIPLICITY (160)

Still essentially pervasive, but with discrete and countable entities, is the group for which **multiplicity** seems the salient feature. Don't bother counting the entities; they are too numerous.

- koṇa* 'comma' *kkōṇaṇa* DR-14 d-f- $\emptyset$   
*Ekkōṇaṇa jaṭōj ɸe.* [3SG- ~ sentence that] 'That sentence has lots of commas'.  
*Ekkōṇaṇa aṇ kōnoṇo.* [3SG- ~ thing-2SG speak] 'You're always pausing in your speech'.<sup>13</sup>
- tiṇoṇ* 'ghost, demon' *ttiṇoṇṇoṇ* DR-14 d-f- $\emptyset$   
*Ettiṇoṇṇoṇ ānin.* [3SG- ~ islet-this] 'This island is haunted'.
- ṇiitwa* 'barracuda' *ṇiitwawa* DR-14  $\emptyset$ -f- $\emptyset$   
*Eṇiitwawa wūnaak in.* [3SG- ~ school.of.fish this] 'There are lots of barracuda in this school of fish'.

Note that this group contains only nominal bases, to which DR-14 applies. Their verbal complements are in the following group.

## FREQUENTATIVE (160)

Verbal bases subjected to DR-18 and DR-19 yield frequent actions.

- raelep* 'to be noon' *rraelelep* DR-18 d-f- $\emptyset$   
*Erraelelep an ruj.* [3SG- ~ thing-3SG waken] 'He usually wakes up at noon'.
- jukwa* 'use sugar' *jjukwakwa* DR-18 d-f- $\emptyset$   
*Kwōmake jjukwakwa.* [2SG-be.alone ~] 'You use sugar too often'.
- po* 'lower sail' *ppopo* DR-18 d-f- $\emptyset$   
*Wa eo eppopo ep.* [canoe the 3SG- ~ that] 'That canoe is stopping at many places. The sail of that canoe keeps coming down'.

<sup>11</sup>We see in this pair of sentences the close connection between the **continuative** and the **pervasive**.

<sup>12</sup>The vowel alternations manifest in these reduplicated forms, and in others that follow, generally conform to regular patterns, which are discussed elsewhere. See, for example, Bender (1973).

<sup>13</sup>Note the metaphorical extension.



<i>tq̄nōt</i>	'make doughnuts'	<i>ttq̄nōtnōt</i>	DR-18	d-f-∅
<i>Lio ettq̄nōtnōt eq̄.</i> [gal-the ~ that] 'She's always making doughnuts'.				
<i>lake</i>	'to lock something'	<i>llaklake</i>	DR-19	d-i-∅
<i>Iṅōk kōn aṅ llaklake.</i> [1SG-be.tired because.of thing-2SG ~] 'I'm tired of you always locking it up'.				
<i>dibōj</i>	'to spear someone/something'	<i>dibdibōj(e)</i>	DR-19	∅-i-(e)
<i>Kwōn jab dibdibōj ek q̄e.</i> [2SG-PTTV not ~ fish that] 'Don't spear so many holes in that fish'.				
<i>Kwōn jab dibdibōje ek q̄e bwe enaaj nana.</i> [2SG-PTTV not ~ fish that because 3SG-IRRL bad] 'Don't spear the fish too many times or it will spoil'.				
<i>Pako eo eq̄ eq̄ōj dibdibōje.</i> [shark the that 3SG-be.finished ~] 'The shark has been speared many times'.				

PARCELLING OUT (20)<sup>14</sup>

Although this group contains the distributives formed from the numerals, it is not limited to them. The emphasis here is on the creation of groups of equal size.

<i>juon</i>	'one'	<i>kajjojo</i> (continuative)		ka-d-f-∅
<i>Rej kajjojo etal.</i> [3PL-PROG ~ go] 'They're going one at a time'.				
<i>Aolep ej kajjojo lq̄k wōt mā.</i> [all 3SG-PROG ~ away only breadfruit] 'Everybody has a breadfruit'.				
<i>jilu</i>	'three'	<i>kajjiluuki</i> (non-continuative transitive)		ka-d-∅-e
<i>Raar kajjiluuki pinana eo nān maat in.</i> [3PL-CMPL ~ banana the to all.gone this] 'They each took three bananas until the bunch ran out'.				
<i>tq̄ujin</i>	'thousand'	<i>kōttq̄ujin</i> (non-continuative intransitive)		ka-d-∅-∅
<i>Rōban kōttq̄ujin.</i> [3PL-be.unable ~] 'They can't each take a thousand'.				
<i>jipañ</i>	'to help someone'	<i>jipjipañ</i>	DR-19	∅-i-∅
<i>Koṅwin itok jen jipjipañ doon.</i> [2PL-PTTV come 1PL-in-PTTV ~ each.other] 'Come let's help each other (and divide up the separate tasks)'.				
<i>pakiji</i>	'to package something'	<i>ppakijkiji</i>	DR-19	d-f-∅
<i>Kwōmarōn ke ppakijkiji pilawā q̄e kijerro?</i> [2SG-be.able Q ~ bread that food-1PL-in-DUAL] 'Can you put our flour in as many packages as necessary?'				

## DISCONTINUOUS (10)

Not only may the entity be poured into separate containers with spaces separating them, the moulds may not all be of the same size or located at regular intervals.

<sup>14</sup>The quantity here, as elsewhere, refers to the approximate number of distributives of this type in our database, not to those theoretically possible.



<i>wiwi</i>	'fat in turtle shell'	<i>uwivi</i>	DR-14	d-0-0
<i>Euwiwi jpkur in.</i> [3SG- ~ turtle.shell this] 'This turtle shell is covered all over the inside with delicious turtle fat'.				
'This turtle shell has fat here and there inside it'. <sup>15</sup>				
<i>mokwaŋ</i>	'pandanus pudding'	<i>mmokwaŋkwaŋ</i>	DR-14	d-f-0
<i>Emmokwaŋkwaŋ nuknuk ŋe aŋ.</i> [3SG- ~ clothes that thing-2SG] 'Your clothes have scraps of pandanus pudding on them'.				
<i>mälle</i>	'charcoal, embers'	<i>mmällele</i>	DR-14	d-f-0
<i>Ej mmällele wöt.</i> [3SG-PROG ~ still] 'There are still some embers there in the ashes'.				
<i>ŋörä</i>	'be dry'	<i>ŋŋöräre</i>	DR-18	d-f-0
<i>Ejino ŋŋöräre nuknuk k̄a aō.</i> [3SG-begin ~ clothes these thing-1SG] 'My clothes are beginning to dry (in certain places)'.				

## DIFFUSE, DILUTE (25)

Nor do we necessarily guarantee that the moulds will be filled pervasively, whatever their size, shape, and location. Some dilution may take place.

<i>arnej</i>	'person, people'	<i>armije</i>	DR-14	0-0-e
<i>Earmije ānin.</i> [3SG- ~ islet-this] 'This island has lots of people'. 'This island is inhabited'. <sup>16</sup>				
<i>marok</i>	'be dark, in darkness'	<i>mmarokrok</i>	DR-18	d-f-0
<i>Ear ruj wöt ke ej mmarokrok.</i> [3SG-CMPL waken still when 3SG-PROG ~] 'He woke as the darkness was vanishing'.				
<i>mijel</i>	'be thick'	<i>mmijeljel</i>	DR-18	d-f-0
<i>Ejaadin mmijeljel wöt.</i> [3SG-be.somewhat ~ only] 'It still is somewhat thickish'.				
<i>tebu</i>	'be fat, overweight, obese'	<i>ttebubu</i>	DR-18	d-f-0
<i>Ejaadin ttebubu.</i> [3SG-be.somewhat ~] 'She's on the chubby side'.				
<i>būrōrō</i>	'be red'	<i>bbūrōrō</i>	DR-18	d-0-0
<i>Ej kōŋak juon nuknuk bbūrōrō mejān.</i> [3SG-PROG wear one clothes ~ face-3SG] 'She is wearing a reddish dress'.				

## AURA, AROMA, AIRS (20)

All that remains is a few tiny pieces or molecules, enough for the odour, or perhaps just the abstract essence. Note that each of these three examples is identified as ambiguous, and that in addition to the meaning that belongs here, each has a more substantive meaning that should be classified elsewhere.

<i>bwiro</i>	'preserved breadfruit'	<i>bbwiroro</i>	DR-14	d-f-0
<i>Ebbwiroro nuknuk e aō.</i> [3SG- ~ clothes this thing-1SG] 'My clothes have scraps of preserved breadfruit on them'. 'My clothes smell of preserved breadfruit'.				

<sup>15</sup>Note the perceived ambiguity on the part of the translator. The first sense should be classified as **pervasive**.

<sup>16</sup>Note another perceived ambiguity. Here again, the first sense belongs under **pervasive**.

- kuwat* 'can of food, tin can' *kkuwatwat* DR-14 d-f- $\emptyset$   
*Eokkuwatwat ꞑwiin*. [3SG- ~ house-this] 'This house is full of cans. This house smells like tin cans'.
- ꞑꞑmaan* 'man, male' *ꞑꞑmaane, ꞑꞑmaanane* (DR-14, 18)  $\emptyset$ - $\emptyset$ -e,  $\emptyset$ -f-e  
*Eꞑꞑ an ꞑꞑmaane ānin*. [3SG-be.large thing-3SG ~ islet-this] 'There are lots of men on this islet'.  
*Eꞑꞑmaanane*. [3SG- ~] 'He's manly'.

## HELTER SKELTER (20)

For distributives of verbal origin, what may be lacking is form or direction.

- jeor* 'to turn' *jjeoreor* DR-18 d-f- $\emptyset$   
*Ejjeoreor wa eo*. [3SG- ~ canoe the] 'The vehicle keeps turning this way and that'.
- jerak* 'hoist sail' *jjerakrōk* DR-18 d-f- $\emptyset$   
*Kōmij jjerakrōk bajjek*. [1PL-E-PROG ~ just] 'We're just sailing around, going no place'.
- kkar* 'arrange something' *kkarrūkarōk, kkarkarōke* DR-19 (irregular)  
*Jab kkarrūkarōkꞑok eok bwe jeeknaan*. [not ~ you because you common] 'Don't try and get yourself accepted (by doing different things) because you're only second class'.  
*Rej kkarkarōke lowaan ꞑweeq*. [3PL-PROJ ~ inside-of house-that] 'They are rearranging the interior of that house this way and that way'.
- jeꞑate* 'take apart, unwind' *jalꞑate* DR-19  $\emptyset$ -i- $\emptyset$   
*Eꞑōj aer jalꞑate injin eo*. [3SG-be.finished thing-3PL ~ engine the] 'They have taken the engine apart carelessly'.
- rupe* 'break, demolish something' *ꞑꞑrupe, ruꞑrupe* DR-19 (d)-i- $\emptyset$   
*Eꞑōj ꞑꞑrupe ꞑweo*. [3SG-be.finished ~ house.the] 'The house has been torn down carelessly (with debris left lying all about)'.
- ebaje* 'scatter' *ebebaj* DR-19  $\emptyset$ -i- $\emptyset$ <sup>17</sup>  
*Eꞑōj an baru ko ebebaj ijo*. [3SG-be.finished thing-3SG crab the-PL ~ there] 'The bulldozers have dug things up there and made the land uneven'.
- medeke* 'to conciliate someone' *mmedekdeke* DR-19 d-f- $\emptyset$   
*Kwōn ilān mmedekdeke jān an ilān kōꞑꞑman tūrabōꞑ*. [2SG-PTTV go-for ~ from thing-3SG go-for make trouble] 'Go do anything you can to keep him from going and making trouble'.
- peꞑote* 'to tear something' *ppeꞑeꞑote* DR-19 d-i- $\emptyset$ <sup>18</sup>  
*Ejjidikdik jān wōt aer kar ppeꞑeꞑote*. [3SG-be.tiny[DSTR] from only thing-3SG PAST ~] 'It was in shreds after they tore it to pieces'.

<sup>17</sup>This seems to be an instance of the omission of the -e transitivity suffix present in the base.

<sup>18</sup>The initial *pe-* of the base in a form such as this is ignored in locating the initial syllable for reduplication. Similarly, in the preceding form, the transitivity suffix -e of the base is ignored in locating the final syllable for reduplication. (The phonemic transcriptions of the two distributive forms are *mmedekdekey* and *ppeꞑawꞑateꞑ*.)

## IDIOMATIC (5)

Finally, we turn to those derived forms for which the meaning connection is not always obvious.

*būri* 'uvula, esophagus' *bbūri* DR-14 d-f-∅  
*Kwōn kōmat lōk bwe ibbūri*. [2SG-PTTV cook away because 1SG-~] 'Hurry up with the cooking because I'm starved'.

*jerakrōk* 'go sailing' *jjerakrōk* DR-18 d-f-∅  
*Ejjerakrōk lōjilnīn*. [3SG - ~ ears-3SG] 'He's got elephant ears'.

*llo* 'see, visit' *llolo* DR-18 d-f-∅  
*Liep ej bar llolo*. [gal-that 3SG-PROG again ~] 'She's unfaithful to her husband'.

## 5.1 SUMMARY OF SEMANTIC GROUPINGS

A rough attempt at charting this semantic constellation artificially in two dimensions might look something like the portrayal in Figure 1.

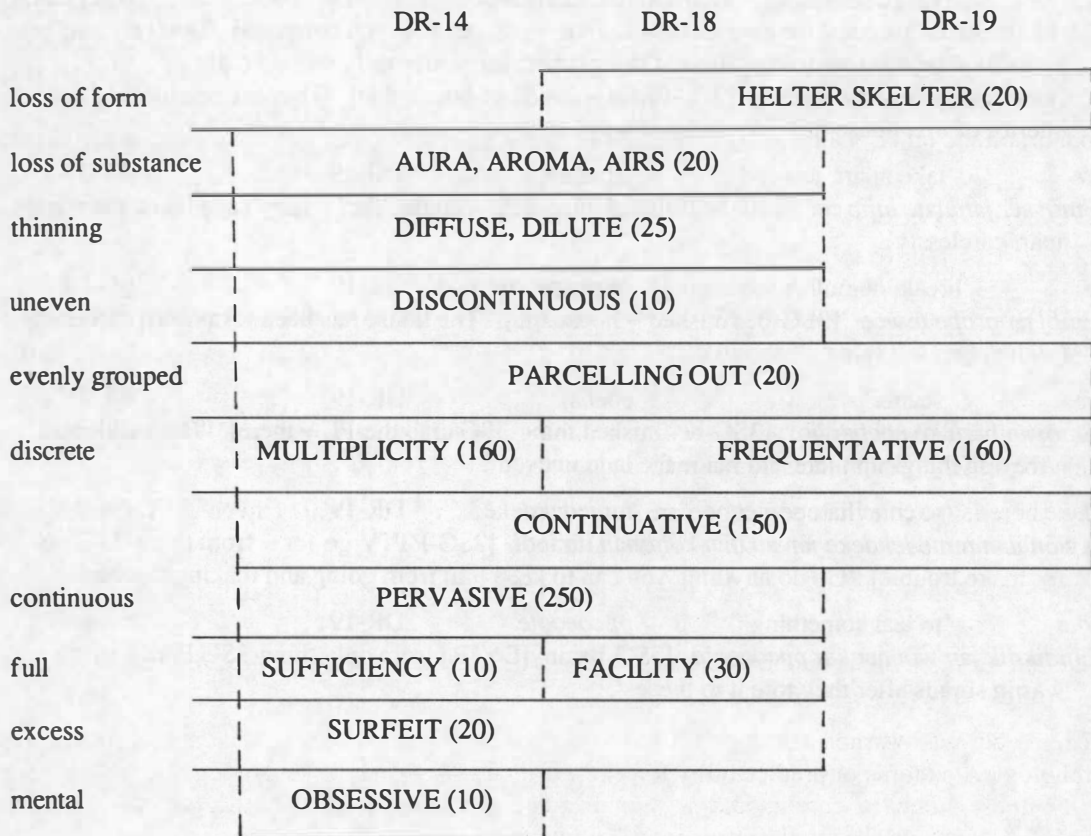


FIGURE 1: SEMANTIC GROUPINGS WITHIN THE CATEGORY 'DISTRIBUTIVE' SHOWN IN RELATION TO EACH OTHER AND TO THE THREE MAJOR DERIVATIONAL RULES THAT PLAY A PART IN THEIR FORMATION.

By using the analogy of a tour in moving from group to group through the examples, as well as in the arrangement of Figure 1, I have emphasised the close interconnections, and perhaps lack of discrete boundaries, between the groups juxtaposed. And although I suggest along the left margin of Figure 1 what may appear to be semantic features, it is my view that the answer as to why a given distributive takes on a particular one of these many shadings will be found more in cognitive grammar<sup>19</sup> than in set theory or any system of distinctive features. By identifying a core area, and at least two separate extensions from it, I have intended to outline a possible basis for such a cognitive approach.

Yet there are indications that we are not dealing here with just an amorphous constellation without any sharp distinctions. Several of the example sentences given have ambiguous readings that have been identified by their translators. These include the distributives based on *wiwi* 'fat in turtle shell' under **discontinuous**, *arnej* 'person, people' under **diffuse**, **dilute**, and *ḡḡmaan* 'man, male' and *kuwat* 'can of food, tin can' under **aura**, **aroma**, **airs**. In each of these instances, the alternative reading has been near the core, either **pervasive** or **multiplicity**. Evidence such as this suggests that there is a basic distinction within the distributive domain related to the two directions of our tour, between a pervasive and concentrated reading, on the one hand, and a scattered, spotty reading on the other. Whether both readings are open to all distributives, or whether one or the other is proscribed for some from the start, will have to remain for future research. Another clear-cut division within the domain may be that between the substantial and the ethereal, as can be seen in many examples that have a double reading between scraps or bits and pieces of something, on the one hand, and simply its odour, on the other. One such is that given for *bwiro* 'preserved breadfruit' under **aura**, **aroma**, **airs**.

## 6. CONCLUSIONS

Much of this paper has dealt with the diversity of distributives in Marshallese, both formal and semantic. Contributing to an overall impression of diversity are:

- (a) the four or five rules necessary for their derivation (including the numerals);
- (b) the fact that, while they are all verbs, they belong to several subclasses of verbs;
- (c) the variety of both their formal marking and the meanings they signify.

Yet there is much evidence that we are dealing with a single semantic category. There are two major indications that point to the unity of the category:

- (a) the way in which the morphological patterns (including the highly productive d-f-∅ pattern) cut across the derivational rules;
- (b) the way in which the narrower semantic rubrics that have been identified also cut across the rules, as portrayed in Figure 1.

The reader was warned at the outset not to expect to find conditioning factors for the various morphological patterns or predictability for the seemingly idiosyncratic semantics. Does this mean that the quest should be continued, that there is some underlying regularity that we have failed to uncover? I think not, for reasons such as the following.

<sup>19</sup>I refer here to recent work of Mark Johnson, George Lakoff and Ronald Langacker, for example.

Distributives in Marshallese, whether inherent or derived, are distinct lexical items, each with a life of its own, independent at least in part from that of any original source. In form, most are distinctively marked. Semantically, they may begin their existence as distributives near the core of that domain. But there can be no guarantee that they will remain there. Aronoff (1976:18) puts it aptly:

But words are peculiar, not only in that not all of those that should exist actually do, but also in that those which do exist do not always mean what they are supposed to mean, or even look like what they are supposed to look like. Words, once formed, persist and change; they take on idiosyncracies, with the result that they are soon no longer generable by a simple algorithm of any generality.

We have taken a close look at an interesting category that has not received a great deal of attention in the literature. We have looked at data from a language in which the category is well developed and plays a prominent role within the lexicon. In many languages it is not morphologically marked, and thus exists only as a semantic option to be dealt with periphrastically. Similar studies are needed from other languages in which it is prominent to help us see how typical the Marshallese picture is. And since there are indications that a number of Austronesian languages are among those in which it is prominent, a knowledge of its distribution within this family, and elsewhere as well, could contribute to our understanding of its evolution as a category.

#### BIBLIOGRAPHY

- ABO, Takaji, Byron W. BENDER, Alfred CAPELLE and Tony DeBRUM**, 1976, *Marshallese-English dictionary*. Honolulu: University Press of Hawaii.
- ARONOFF, Mark**, 1976, *Word formation in generative grammar*. Cambridge: The MIT Press.
- BENDER, Byron W.**, 1969, *Spoken Marshallese: an intensive language course with grammatical notes and glossary*. Honolulu: University of Hawaii Press.
- 1973, Parallelisms in the morphophonemics of several Micronesian languages. *Oceanic Linguistics* 12:455-477.
- 1984, Object marking in Marshallese. In Byron W. Bender, ed. *Studies in Micronesian Linguistics, PL, C-80*, 443-465.
- BLAKE, Frank R.**, 1925, *A grammar of the Tagalog language*. American Oriental Series 1. New Haven: American Oriental Society.
- De la CRUZ, Beato A. and R. David Paul ZORC**, 1968, *A study of the Alkanon dialect*, vol. one: grammar. Kalibo, Aklan: Public Domain.
- GREENOUGH, J.B., G. L. KITTREDGE, A.A. HOWARD and Benj. L. D'OUGE**, 1903, *Allen & Greenough's new Latin grammar*. Exact reprint in College Classical Series, New Rochelle, New York: Caratzas, 1983.
- LAKOFF, George**, 1987, *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: The University of Chicago Press.
- LAKOFF, George and Mark JOHNSON**, 1980, *Metaphors we live by*. Chicago: The University of Chicago Press.
- LANGACKER, Ronald**, 1986, *Foundations of cognitive grammar*, vol 1. Stanford: Stanford University Press.
- PAGOTTO, Louise**, 1987, Verb subcategorization and verb derivation in Marshallese: a lexicase analysis. Doctoral dissertation, University of Hawaii, Honolulu.



## SOUND CHANGE AND MIGRATION DISTANCE

ROBERT BLUST

### 1. INTRODUCTION

This paper examines two variables in historical linguistics – amount of sound change and migration distance – and questions whether a significant correlation can be established between them. The original inspiration for the thesis explored here derives from the often-repeated classroom joke that the Polynesians must have ‘tossed phonemes’ overboard as they pressed ever eastwards, Hawaiian (with eight consonants and five vowels plus length) possessing the second smallest phoneme inventory of any known language. Although this observation is based on relative degrees of phonological merger, various observations about shift and split in other Austronesian languages can be taken impressionistically to support the same general idea, namely, that the amount of sound change in a language is positively correlated with the distance its speakers have travelled from the probable homeland of the immediate or wider genetic group to which it belongs.

In addition to these impressions of my own, two other Austronesian linguists in recent years have either implicitly or explicitly maintained that there is a general relationship between sound change and migration distance. Dahl (1976) attributed to Otto Dempwolff the view that the Austronesian homeland could not have been in the region in which the modern Oceanic languages are spoken, since the Oceanic languages exhibit far greater phonological erosion than the typical Austronesian language of island Southeast Asia.<sup>1</sup> Clearly implied in Dahl's interpretation of Dempwolff is a general theoretical claim that languages which are further removed from the primary centre of dispersal will be phonologically more innovative.

During a brief conversation at the Fifth International Conference on Austronesian Linguistics in Auckland, New Zealand in January 1988, Malcolm Ross and I exchanged views of a similar kind. He felt, from his intensive comparative studies of the Austronesian languages of western Melanesia, and I felt, from my knowledge of Austronesian as a whole, that a global relationship between amount of sound change and migration distance was likely. Ross has now published his own views on the matter (this volume). Ross's study is an example of what Matisoff (1990) would call a ‘microcomparison’: it attempts to consider in detail all of the identifiable sound changes for a small group of closely related languages with a separation time probably not much in excess of 1,500 years. He concludes that there is a clear “correlation between sedentariness and conservatism”.

---

<sup>1</sup>See Blust (1976:233) for a critique of this reading of the passage in question.

In this paper I attempt to investigate, through macrocomparison, the same issue addressed by Ross. As will be seen, it turns out to be worthwhile to approach a similar question through both micro- and macrocomparison, as the explanatory mechanisms available for one may not always be available for the other. As a test case I examine the 'erosion sequence'  $*p > f > h > \text{zero}$  and certain variants of this development in Austronesian languages. Proto Austronesian canonical shape was CV(C)CVC, where (C) was either the abutting consonant in a reduplicated monosyllable, or a nasal generally homorganic with a following obstruent. I have confined the present study to reflexes of initial and intervocalic  $*p$ , since to have done otherwise would have greatly complicated the evaluation of the results.

My views on the Austronesian homeland and the major migration routes leading from island Southeast Asia into the Pacific have been stated elsewhere (Blust 1984/85). They can be summarised as follows: 1) Proto Austronesian was spoken on or near Taiwan, 2) the first migration from this centre, sometime before 4000 B.C., was southwards into the Philippines, 3) from the southern Philippines the stream of migration split, one branch (ancestral to the Western Malayo-Polynesian (WMP) languages) moving into Borneo and Sulawesi, the second branch (ancestral to the Central-Eastern Malayo-Polynesian (CEMP) languages) moving into the northern Moluccas, 4) from the northern Moluccas the stream of migration split again, one branch (ancestral to the Central Malayo-Polynesian (CMP) languages) moving south and west through the Moluccas and Lesser Sunda Islands as far as Sumbawa, the second branch (ancestral to the Eastern Malayo-Polynesian languages) moving into west New Guinea, 5) Proto Eastern Malayo-Polynesian split into Proto South Halmahera-West New Guinea and Proto Oceanic; some of the descendants of the former moved back westwards into the northern Moluccas, while the descendants of the latter continued to press south and east to the north coast of New Guinea facing the Bismarck Archipelago, 6) from the New Guinea-Bismarck Archipelago region, a rapid dispersal eastwards into the Solomon and Vanuatu chains, southern Melanesia, Micronesia and the central Pacific completed the expansion of the Eastern Malayo-Polynesian-speaking peoples, 7) from Borneo and Sulawesi the further movement of Western Malayo-Polynesian-speaking peoples was more restricted, the most spectacular exception being the settlement of Madagascar from south-east Borneo in the early centuries of the Christian era (Dahl 1951).

In discussing spatial distribution I have divided the languages into 16 major geographical groupings, displayed in Table 1. The smallest of these geographical groupings (Madagascar) contains a single language, the largest (New Guinea) contains 161. The relevance of subgrouping to the matter at hand will be discussed in subsequent sections.

Section 3 presents a summary of the data. In order to process any of the data, however, a number of methodological issues must be addressed, and this is undertaken in section 2.

## 2. THE QUANTIFICATION OF PHONOLOGICAL CHANGE

Before we can correlate the amount of sound change with migration distance, we must quantify sound change cross-linguistically. To attempt this for all sound changes in even a small language family would be a major undertaking; to do it for a language family such as Austronesian, with over 900 member languages, would require several volumes. Given the practical limitations of my paper I have chosen to represent sound change as a whole by a series of changes that I will call the 'voiceless bilabial erosion sequence':  $*p > f > h > \text{zero}$  (as qualified below). The general validity of my thesis, at least as regards this particular development, has long been implicitly recognised by Austronesian



linguists. So rarely is *\*p* altered in Philippine languages that a pioneering Philippinist devoted a paper especially to the changes *\*p > f* and *\*b > v* (Conant 1908). To anyone familiar with the historical phonology of Central Malayo-Polynesian or Eastern Malayo-Polynesian languages such an enterprise would appear strangely out of place, since *\*p* in the languages of eastern Indonesia and Oceania is rarely retained as a stop.<sup>2</sup>

## 2.1 PARAMETER SETTING

For purposes of quantification I have set the numerical parameters of the erosion sequence *\*p > f > h > zero* at 0, 1, 2, 3 respectively. In other words, each language which shows no change of *\*p* will be assigned a numerical value of zero, each language which shows a change of *\*p* to [f] will be assigned a numerical value of one, and so on. Since the relative proportions are unchanged, it is irrelevant whether the parameters are set at 0, 1, 2, 3, at 1, 2, 3, 4 or at any other arbitrary range of numerical values with constant intervals.

To illustrate, imagine a hypothetical set of 30 languages with the following distributions on the *\*p* 'erosion sequence': *\*p > [p]* (one case), *\*p > [f]* (eight cases), *\*p > [h]* (eleven cases), *\*p > zero* (ten cases). Multiply the first case (zero) by one, the second case (1) by eight, the third case (2) by eleven and the fourth case (3) by ten, giving 60. Dividing this figure by the total number of languages for which erosion values are computed (30) gives 2.0 (corresponding to [h]) as the mean erosion value for the set of languages compared.

Since the diachronic attribute that I am attempting to measure is a phonetic characteristic, I have used phonetic values wherever they differ from phonemic values, for purposes of calculating degree of erosion. Similarly, I make the relatively non-controversial assumption that Proto Austronesian (PAN) *\*p* was [p].

## 2.2 HETEROGENEOUS REFLEXES

Many sound changes are conditioned (e.g. Tagabili, where *\*p > h-*, *-f-*), while others show unexplained splits of a proto-phoneme (e.g. Miri, where *\*p > /f/* or */p/* without known conditions). In all such cases I have averaged the erosion values for the various reflexes without consideration of relative frequency. Thus, for Tagabili of the southern Philippines the erosion value is 1.5 (mid-way between [f] and [h]), while for Miri of western Borneo it is 0.5 (mid-way between [p] and [f]).

Special complications arise in the case of the Oceanic languages. In Proto Oceanic, PAN *\*b* and *\*p* merged as *\*p* (often reflected as a fricative) and PAN *\*mb* and *\*mp* merged as *\*mp* (often reflected as a bilabial stop). A certain amount of unpredictable 'cross-over' of consonant grade has long been reported for Oceanic languages (the expected reflex of POC *\*p* in a given morpheme actually being the reflex of *\*mp*, or vice versa). As shown by Ross (1988), many Oceanic languages have two reflexes of *\*p* that are independent of this phenomenon of oral/nasal 'grades'. He refers to these as 'fortis' and 'lenis' reflexes. Wherever the 'fortis' reflex of *\*p* is identical to the nasal grade reflex I treat it as an instance of the latter; where it differs from the nasal grade reflex I treat it as a case of unexplained phonemic split, and average the numerical values of the reflexes, as with the Miri case noted above.

<sup>2</sup>Although this study is inherently statistical, I have made no use of statistical tests of significance, which often serve as little more than a confirmation of the obvious. As will be seen, the numerical values upon which my conclusions rest leave little doubt about the significance of the groupings obtained.

## 2.3 SEQUENCE EQUIVALENTS

Proto Austronesian *\*p* has 17 known reflexes: 1) *b*, 2) *β*, 3) *ð*, 4) *f*, 5) *h*, 6) *k*, 7) *p*, 8) *ɸ* (apico-labial stop), 9) *pw*, 10) *ʔ*, 11) *t*, 12) *θ*, 13) *v*, 14) *ʋ* (apico-labial fricative), 15) *w*, 16) *y* and 17) zero. Some of these reflexes are quite rare or are found only in highly restricted environments. For our purposes what is important is that the values [p], [f], [h] and zero form a well-established diachronic sequence involving increasing lenition (weakening of articulatory stricture).<sup>3</sup> Since [v] and [w] are also quite common reflexes, the question naturally arises whether either of these segments can be placed in the established erosion sequence. For purposes of the present paper I have disregarded voicing as a factor in the evaluation of degrees of lenition. As a result an equivalence is established between [f] and [v] and between [h] and [w]: parallel to the main erosion sequence sketched above, a number of Oceanic languages have participated in an equivalent sequence *\*p* > [f] > [v] > [w]. The numerical value for [v] as a reflex of *\*p* is thus set at (1), and the numerical value of [w] as a reflex of *\*p* is set at (2).

## 2.4 SECONDARY FORTITION

The adoption of an 'erosion sequence' model for the historical development of PAN *\*p* presupposes a unilateral direction of change. However, in a small number of cases it appears that a change *\*p* > [f] has reversed itself, thereby giving the prima facie impression that no change has occurred. This is seen in the Polynesian Outlier of Anuta, in which PPN *\*p* and *\*f* have merged as a voiceless bilabial stop (Feinberg 1977), and in at least two Micronesian languages, Ponapean and Mokilese, in which PMC *\*p* and *\*f* have merged as [p] (Bender, et al. 1984). For purposes of determining erosion values I have taken these reflexes at face value, although it is likely that they embody a complex history leading from stop to fricative and back to stop.

Needless to say, it is possible that some other instances of *\*p* > [p] are also products of secondary fortition. This is most probable in linguistic subgroups where the reflex *\*p* > [p] is rare.

## 2.5 ATOMISM

The material summarised in section 3 is taken from sources which do not pretend to describe all dialectal variations, and it is possible that some languages for which a given erosion value of *\*p* is reported have dialects with different erosion values. To have pursued the study to this level, however, would have greatly increased the number of reflexes that require checking, and in all probability would have affected the results very little.

Under certain imaginable circumstances the inclusion of dialect material could significantly affect mean erosion value. For example Malagasy, which is normally represented by the Merina dialect, is a dialect complex with at least 18 distinguishable varieties (Vérin, Kottak and Gorlin 1969). All Malagasy dialects reflect *\*p* as [f], and the mean erosion value for Madagascar thus remains 1.0 whether we count only Merina, or all 18 recognised dialects. Malagasy subgroups with the languages of Borneo, and if it were included in that geographical region the mean erosion value for the languages of Borneo would differ considerably depending on whether Malagasy dialects were

<sup>3</sup>Little work has yet been done on diachronic universals, but preliminary investigation of other language families suggests that this erosion sequence is universal both in the sense that it is commonly attested, and in the sense that, with rare exceptions (e.g. *\*f* > [p]), it always reads from left to right. This is true regardless of the source of a non-initial sequence member (e.g. Latin [f], from the merger of PIE *\*bh*, *\*dh* became written Spanish /h/, spoken Spanish zero).

counted separately or not. But since our purpose is to examine possible correlations between amount of phonological change and migration distance, it is clear that Malagasy must be treated as a geographical region separate from Borneo, regardless of its subgrouping affiliations. To my knowledge, despite the potential for such a situation to exist, no actual dialect situation would significantly affect the mean erosion value of a major geographical region if we were to adopt a more atomistic approach to the data.

## 2.6 GALTON'S PROBLEM

Galton's problem (named after Francis Galton, who first clearly stated it a century ago in public discussion with E.B. Tylor) formulates a basic condition on comparative research in the social sciences. If the sample units in a comparative corpus are not truly independent the results of the comparison will show sampling bias. A good example is the attempt of Lees (1953) to establish universal replacement rates of basic vocabulary through a pilot study which involved 13 languages, of which 11 are Indo-European, and six of these Romance. In general, any attempt to establish language universals inductively is subject to the same criticism.

In the present case our decision to treat each language as an independent sample unit is a fiction, but one which cannot be avoided until a complete subgrouping, down to the lowest levels of inclusion, is available for the entire Austronesian language family. To take a particularly clear example, a subgroup including all and only the Polynesian languages is universally accepted, and it is universally agreed that in its immediate ancestor PAN *\*p* had become *\*f*. To count each instance of PAN *\*p* > [f] in a Polynesian language as an independent example of phonological erosion is simply to multiply a single case by the number of its descendants. In the present study I have not been able to control for Galton's problem, but despite this limitation the established correlations appear to be generally valid, since many languages which have inherited one step in the erosion of *\*p* from a common subgroup ancestor have subsequently undergone further changes on their own.

## 3. DATA SUMMARY AND DISCUSSION

Data was obtained by scanning the available sources for 930 Austronesian languages, which are grouped by major geographical areas. Table 1 presents a summary of the results concerning the general lenition of *\*p*, without reference to specific erosion values: (1) = number of languages, (2) = percentage of Austronesian total, (3) = *\*p* unchanged, (4) = *\*p* changed, (5) = development of *\*p* unknown, (6) = percentage of languages with some change of *\*p* (rounded to nearest whole figure).

TABLE 1: A GLOBAL SUMMARY OF THE LENITION OF PAN *\*p* RELATED TO MAJOR GEOGRAPHICAL AREAS

No.	AREA	(1)	(2)	(3)	(4)	(5)	(6)
1.	Taiwan	22	2.4	18	2	2	10
2.	Philippines	129	13.9	69	14	46	17
3.	Borneo	96	10.3	62	2	32	3
4.	Madagascar	1	.1	0	1	0	100
5.	Mainland SEA	10	1.1	10	0	0	0
6.	Sumatra-Sumbawa	21	2.3	16	2	3	11
7.	Sulawesi	67	7.2	34	4	29	11

8. Lesser Sundas	46	4.9	18	18	10	50
9. Moluccas	53	5.7	2	36	15	95
10. New Guinea	161	17.3	2	133	26	99
11. Bismarck Archipelago	74	7.9	11	56	7	84
12. Solomons-Santa Cruz	70	7.5	0	68	2	100
13. Micronesia	24	2.6	0	20	4	100
14. Vanuatu	105	11.3	0	91	14	100
15. New Caledonia	28	3.0	0	11	17	100
16. Rotuma, Fiji, Polynesia	23	2.5	0	22	1	100
	930	100	242	480	208	

For 208 of these languages information was insufficient to determine the reflex of initial and intervocalic \*p. The remaining 722 languages are divided between two categories: (1) those in which \*p > [p] is the *only* development, and (2) those in which a change of \*p is *at least one* of the known developments. A brief discussion of each geographical area follows.<sup>4</sup>

#### TAIWAN

For Basai (sometimes called Ketangalan) and Ketangalan (sometimes called Luilang) the available information is insufficient to determine reflexes (Tsuchida 1982, 1985). The two Formosan languages in which \*p clearly has changed are Kulon, where \*p > [p]/[h] (Tsuchida 1985) and Favorlang, where \*p > [p]/[h]/zero (Marsh 1977). In neither case can conditions for the split be stated. The erosion values (EV) in these two cases are 1.0 and 1.66. The mean erosion value for all Formosan languages for which adequate information is available is thus  $(18 \times 0) + (1 \times 1.0) + (1 \times 1.66) = 2.66/20 = .133$ .

#### PHILIPPINES

According to Conant (1908), in Ibanag and the closely related Gaddang, Yogad and Itawis of northern Luzon, \*p became [f] only before \*u. In early twentieth-century Ibanag this change had begun to generalise to other environments, although not all relevant lexical items were yet affected. In Ibaloi of north-central Luzon \*p > [f] is said to be a sporadic development. The erosion value in each of these five cases is 0.5.

Zorc (1974) has provided data on Buhid of Mindoro, in which \*p appears to have shifted unconditionally to [f]. This is the only reported example of the lenition of \*p anywhere in the central Philippines.

In Tiruray, Bilaan, Blit Manobo and Tasaday of southern Mindanao, \*p is unconditionally reflected as [f], while in Tagabili \*p is reflected as [h] word-initially and as [f] intervocalically (EV = 1.5). In Kalagan (Tagakaolo), Sarangani Manobo and Cotabato Manobo \*p is reflected as [ϕ], here treated as equivalent to [f] (EV = 1.0). The mean erosion value for all Philippine languages is thus  $(69 \times 0) + (5 \times 0.5) + (1 \times 1.5) + (8 \times 1.0) = 12/83 = .145$ .

<sup>4</sup>The language list is a slightly modified version of Wurm and Hattori (1981). Where not otherwise noted major sources of data are as follows: 1) Taiwan: Ferrell (1969), Tsuchida (1971), 2) Philippines: Reid (1971), McFarland (1977), 3) Borneo: Ray (1913), Hudson (1967), Prentice (n.d.), Blust (n.d.), 4) Mainland Southeast Asia: (Lee 1966), Benedict (1984), 5) Sulawesi: Sneddon (1978), (1984), Himmelmann (1990), Mills (1975), van den Berg (1988), 6) Lesser Sundas: Fox (n.d.), 7) Moluccas: Stresemann (1927), Collins (1982), (1983), 8) New Guinea: Ross (1988), 9) Bismarck Archipelago: Ross (1988), Blust (n.d.), 10) Solomons-Santa Cruz: Tryon and Hackman (1983), 11) Micronesia: Bender et al. (1984), 12) Vanuatu: Tryon (1976), 13) New Caledonia-Loyalties: Haudricourt (1971), Haudricourt and Ozanne-Rivierre (1982), 14) Rotuma-Fiji-Polynesia: Biggs (1978).

## BORNEO

Banggi, spoken on a small island off the north coast of Sabah, shows  $*p > [f]$  (EV = 1.0), and Miri, spoken near the mouth of the Baram River in northern Sarawak, shows  $*p > [p]/[f]$  (EV = 0.5). The mean erosion value for all languages of Borneo is thus  $(62 \times 0) + (1 \times 1.0) + (1 \times 0.5) = 1.5/64 = .023$ .

## MADAGASCAR

A single Austronesian language, Malagasy, is spoken on Madagascar. In Malagasy initial and intervocalic  $*p$  have become  $[f]$ . The mean erosion value for Madagascar is thus 1.0

## MAINLAND SOUTHEAST ASIA

The Austronesian languages of mainland Southeast Asia include Malay, Moken and the eight members of the Chamic group. No change has affected  $*p$  in any of these languages. The mean erosion value of this geographically-defined linguistic unit is thus zero.

## SUMATRA-SUMBAWA

For the geographical area including Sumatra, Java, Madura, Bali, Lombok and western Sumbawa only two languages are known to show a change of  $*p$ : Simalur (Simeuleu), where initial  $*p$  disappeared and intervocalic  $*p$  became  $[h]$  (EV = 2.5), and Nias, where initial and intervocalic  $*p$  became  $[f]$  before all vowels (EV = 1.0). The mean erosion value for this geographical region is thus  $(16 \times 0) + (1 \times 2.5) + (1 \times 1.0) = 3.5/18 = .194$ .

## SULAWESI

Several of the languages of extreme south-eastern Sulawesi show lenition of  $*p$ : Muna and Busoa, where  $*p > [f]$  (van den Berg 1988), Tukangbesi, where  $*p$  is reflected as  $[p]$  or zero (EV = 1.5), and Bonerate, where  $*p$  is reflected as  $[p]$  or  $[h]$  (EV = 1.0). The mean erosion value for the languages of Sulawesi is thus  $(34 \times 0) + (2 \times 1.0) + (1 \times 1.5) + (1 \times 1.0) = 4.5/38 = .118$ .

## LESSER SUNDAS

When we reach the languages of the Lesser Sundas the mean erosion value of  $*p$  increases dramatically. Although a number of languages, including Komodo, the languages of Sumba, the languages of Flores, Kédang and such Timorese languages as Kemak and Tukudede reflect  $*p$  as  $[p]$ , fully half of the languages for which information is to hand show some change. Reflexes and erosion values are as follows:

EV = 0.5: 1)  $[p]/[f]$  (Bimanese);

EV = 1.5: 1)  $[p]/\text{zero}$  (Savu), 2)  $[f]/[h]$  (Mambai);

EV = 2.0: 1)  $[f]/\text{zero}$  (Yamdena), 2)  $[h]$  (Galoli, Waima'a);

EV = 2.5: 1)  $[h]/\text{zero}$  (Roti, Helong, Atoni, Tetun, Idate);

EV = 3.0: (Ndao, Wetar, Kisar, Roma, Leti-Moa, Masela-South Babar, Selaru).

The mean erosion value of  $*p$  in languages of the Lesser Sundas is thus  $(18 \times 0) + (1 \times 0.5) + (2 \times 1.5) + (3 \times 2.0) + (5 \times 2.5) + (7 \times 3.0) = 43/36 = 1.194$ .

## MOLUCCAS

Only two Moluccan languages, Teor-Kur and Buru, are known to reflect initial and intervocalic  $*p$  invariably as  $[p]$ . Other erosion values are as follows:



EV = 1.0: 1) [p]/[h] (Soboyo), 2) [f] (Kola, Kei-Fordata, Watubela, Geser-Goram, Masiwang, Paulohi and Ambelau of the southern and central Moluccas, Onin of the Bomberai Peninsula, and Weda, Maba and Buli of southern Halmahera), 3) [ɸ] (Ujir, Dobel, Barakai and Wokam-Tarangan of the Aru Islands);

EV = 2.0: 1) [h] (Bobot/Hatumeten, Manusela, Amahai, Wemale, Loun, Nusalaut, Saparua, Haruku, Kayeli, Giman and East Makian), 2) [ʔ] (treated as equivalent in erosion value to [h]) (Nuaulu, Kaibobo and Asilulu);

EV = 2.5: 1) [h]/zero (Kamarian), 2) [h]/[ʔ]/zero (Hitu);

EV = 3.0: (Banda, Alune, Larike-Wakasihu and Batumerah).

The mean erosion value of \**p* in languages of the Moluccas is thus  $(2 \times 0) + (16 \times 1.0) + (14 \times 2.0) + (2 \times 2.5) + (4 \times 3.0) = 61/38 = 1.605$ .

#### NEW GUINEA

Most of the Austronesian languages of New Guinea are divided between the South Halmahera-West New Guinea and Oceanic subgroups, with a small number of Central Malayo-Polynesian outliers in south-west Irian (Sekar, Uruangnirin, Kaiwai). Despite this relative diversity the languages of this area almost universally show some erosion of \**p*. Dusner (SHWNG) and Tarpia (Oceanic) reflect \**p* as [p], but all other languages for which data are available show erosion in at least one position. The known developments are summarised in the following paragraphs:

EV = 0.5: 1) [p]/[v] (Gitua, Piu, Kapin, Mumeng, Mapos Buang, Manga Buang, Vehes, Hote, Yamap, Misim, Kaiwa, Dawawa, Igora, Nara, Molima, Bosilewa, Kurada and Muyuw), 2) [p]/[f] (Yalu, Sirak, Wampar, Sirasira, Sukurum, Silisili, Dangkal, Maralango, Wagawaga and Mekeo), 3) [b]/[f] (Kuni);

EV = 0.66: 1) [p]/[v]/[ɣ] (Garuwahi);

EV = 1.0: 1) [f] (Maya, Sekar, Uruangnirin, Kaiwai, Biak, Ron, Irahutu, Sobei, Bongo and Adzera), 2) [f]/[v] (Wogeo and Fagululu), 3) [v] (Bwaidoka), 4) [p]/[h] (Musom, Mari, Wampur, Motu and Doura), 5) [p]/[w] (Malamalai, Mutu and Kehelala), 6) [p]/[b]/[h] (Doga, Mukawa, Gapapaiwa, Boianaki and Roro), 7) [p]/[v]/[h] (Bunama), 8) [p]/[b]/[w]/[h] (Bukawac);

EV = 1.25: 1) [p]/[b]/[w]/zero (Anuki), 2) [p]/[v]/[ɣ]/zero (Kukuya);

EV = 1.33: 1) [f]/[v]/[w] (Kairiru), 2) [p]/[f]/zero (Takia, Gedaged, Mindiri, Biliau, Wab and Bohutu), 3) [p]/[v]/zero (Wedau, Taupota, Bina, Magori, Yoba, Ouma, Keapara, Gabadi, Misima and Nimowa), 4) [b]/[f]/zero (Arifama-Miniafia and Ubir), 5) [b]/[v]/zero (Sud-est), 6) [p]/[w]/[h] (Duau);

EV = 1.5: 1) [p]/zero (Ali, Ulau-Suain, Lukep, Roindji, Malasanga, Sewa Bay, Barim, Guwot, Mangap-Mbula, Sio, Tami, Kela, Dobu), 2) [p]/[f]/[h]/zero (Bilbil, Matukar), 3) [f]/[v]/[ɣ]/zero (Sinagoro, Kalokalo, Yamalele), 4) [p]/[v]/[w]/zero (Kilivila, Budibud);

EV = 1.66: 1) [p]/[w]/zero (Kis, Bam, Manam, Medebur, Yabim, Numbami, Tubetube);

EV = 1.75: 1) [p]/[h]/[w]/zero (Suau), 2) [f]/[v]/[w]/zero (Iduna, Diodio);

EV = 2.0: 1) [h] (Mor, Labu), 2) [v]/zero (Gumasi);

EV = 2.5: 1) [h]/zero (Ham, Nengaya), 2) [w]/zero (Kaiep);



EV = 3.0: (Wandamen, Waropen, Ansum, Woi, Pom, Marau, Munggui, Papuma, Busami, Serui-Laut, Ambai, Wadapi-Laut, Kurudu, Sera, Sissano, Tumleo).

The mean erosion value for reflexes of PAN \*p in the Austronesian languages of New Guinea is thus  $(2 \times 0) + (29 \times 0.5) + (1 \times 0.66) + (28 \times 1.0) + (2 \times 1.25) + (21 \times 1.33) + (20 \times 1.5) + (7 \times 1.66) + (3 \times 1.75) + (3 \times 2.0) + (3 \times 2.5) + (16 \times 3.0) = 182.08/135 = 1.348$ .

#### BISMARCK ARCHIPELAGO

The prima facie evidence suggests that the Austronesian languages of the Bismarck Archipelago – particularly those of New Britain – are relatively conservative in their erosion values for \*p. Eleven languages (Mengen, Mamusi, Mangseng, Kapore, Pasismanua, Lamogai, Mok-Aria, Arawe, Longa, Kilenge, Maleu) reflect \*p as [p], and a twelfth (Okro) reflects \*p as [p]/[b]. Other erosion values are:

EV = 0.5: 1) [p]/[f] (Wuvulu-Aua, Ponam, Kara, Konomala), 2) [p]/[β] (Andra-Hus), 3) [p]/[v] (Tabar, Tolai, Melamela, Nakanai, Xarus, Bola, Bulu, Bali-Vitu);

EV = 1.0: 1) [f] (Kaniet, Nalik, Tanga), 2) [p]/[h] (Bipi-Sisi, Likum, Levei-Tulu, Pelipowai, Ere-Lele-Gele-Kuruti, Leipon, Titan, Nali, Loni, Mokerang, Pak-Tong, Lenkau, Penchal, Nauna, Patpatar), 3) [b]/[h] (Hermit, Lindrou, Sori-Harengan);

EV = 1.33: 1) [p]/[f]/zero (Siar), 2) [p]/[v]/zero (Kandas, Duke of York);

EV = 1.5: 1) [p]/zero (Baluan-Lou-Pam, Mussau-Emira, Lavongai, Tigak, Tiang, Notsi, Lavatbura-Lamusong, Madak, Barok, Kaliai-Kove, Bariai);

EV = 1.66: 1) [p]/[h]/zero (Papitalai, Lihir, Sursurunga);

EV = 2.0: 1) [h] (Seimat, Uvol), 2) [v]/[h]/zero (Tomoip);

EV = 3.0: (Tenis).

The mean erosion value for PAN \*p in languages of the Bismarck Archipelago is thus  $(12 \times 0) + (13 \times 0.5) + (21 \times 1.0) + (3 \times 1.33) + (11 \times 1.5) + (3 \times 1.66) + (3 \times 2.0) + (1 \times 3.0) = 62/67 = .925$ .

#### SOLOMON AND SANTA CRUZ ISLANDS

From the Solomon Islands eastwards the oral grade of PAN \*p is never exclusively [p], although [p] may be one reflex.<sup>5</sup> Erosion values are as follows:

EV = 0.5: 1) [p]/[v] (Hahon, Teop, Piva, Papapana, Vaghua, Varisi, Ghanongga, Lungga, Simbo, Nduke, Roviana, Kusaghe, Hoava, Ughele, Marovo, Vanguu, Talise, Lengo, Birao, Longgu), 2) [b]/[v] (Babatana);

EV = 0.66: 1) [p]/[v]/[ɣ] (Banoni);

EV = 1.0: 1) [f] (Laghu, Kokota, Zazao, Blablanga, Gao, Lau, Langalanga, North Malaita, Kwaio, Dori'o, Tikopia), 2) [v] (Bugotu, Nggela, West Guadalcanal, Malango), 3) [f]/[v] (Kia), 4) [p]/[h] (Halia, Solos, Petats, Mono-Alu) 5) [b]/[f]/[h] (Cheke Holo) 6) [p]/[v]/[w] (Tanimbili);

EV = 1.33: 1) [p]/[v]/zero (Timputz, Ririo, Uruava);

EV = 1.5: 1) [f]/[h] (Kwara'ae, Takuu), 2) [v]/[h] (Bauro, Kahua), 3) [p]/zero (Torau);

<sup>5</sup>The Polynesian Outlier language of Anuta in the Solomons chain reflects PAN \*p invariably as [p]. However, as noted in the discussion of 'secondary fortition' (section 2.4), Anuta [p] reflects Proto Polynesian \*f.

EV = 1.66: 1) [p]/[w]/zero (Nehan);

EV = 2.0: 1) [h] (Marau, 'Are'are, Sa'a-Ulawa, Oroha, Arosi, Fagani, Rennell-Bellona, Luangiua, Sikaiana, Pileni), 2) [v]/zero (Nembaou, Asumboa, Vano, Buma, Tanima).

The mean erosion value for PAN \**p* in languages of the Solomon and Santa Cruz Archipelagoes is thus  $(21 \times 0.5) + (1 \times 0.66) + (22 \times 1.0) + (3 \times 1.33) + (5 \times 1.5) + (1 \times 1.66) + (15 \times 2.0) = 76.32/68 = 1.122$ .

#### MICRONESIA

The languages of Micronesia are genetically diverse. Most belong to the Oceanic subgroup of Austronesian, but Palauan and Chamorro are generally classified as co-ordinate branches of Western Malayo-Polynesian, and the position of Yapese remains enigmatic. Apart from Ponapean and Mokilese, which appear to show secondary fortition of Proto Micronesian \**f*, all of the languages of Micronesia, both Oceanic and non-Oceanic, show some erosion of PAN \**p*. Erosion values are as follows:

EV = 1.0: 1) [f] (Mortlockese, Trukese, Puluwat, Satawalese, Woleaian, Chamorro, Saipan Carolinian, Ulithian, Yapese, Sonsorolese);

EV = 1.5: 1) [p]/zero (Ponapean, Mokilese);

EV = 2.0: 1) [y] (Marshallese), 2) [h] (Nukuria, Kapingamarangi, Nukuoro), [w] (Palauan);

EV = 3.0: 1) (Nauru, Gilbertese, Kosraean);

The mean erosion value for PAN \**p* in the languages of Micronesia is thus  $(10 \times 1.0) + (2 \times 1.5) + (5 \times 2.0) + (3 \times 3.0) = 32/20 = 1.600$ .

#### VANUATU

What appears to be the oral grade of \**p* is retained as a stop in a few of the languages of Vanuatu, but only in certain positions. Erosion values are as follows:

EV = 0.5: 1) [p]/[v] (Morouas, Butmas-Tur, Akei, Fortsenal, Polonombauk, Litzlitz, Repanbitip, Malfaxal, Port Vato);

EV = 1.0: 1) [v] (Vunapu, Piamatsina, Tolomako, Malmariv, Navut, Lametin, Amblong, Wailapa, Narango, Tutuba, Malo, North-east Aoban, Central Maewo, Baetora, Vovo, Malua Bay, Mae, Larevat, Vinmavis, Lingarak, Katbol, Dixon Reef, Unua, Rerep, Letemboi, Aulua, Burmbar, Port Sandwich, Axamb, Maskelynes, Ura, Sie), 2) [f] (Mele-Fila, Futuna-Aniwa), 3) [v]/[θ] (Lorediakarkar, Shark Bay), 4) [v]/[ð] (Roria, Tambotalo), 5) [v]/[v̥] (Tangoa, Mafea, Aore, Vao), 6) [p]/[v]/[w] (Apma, Uripiv-Wala-Rano);

EV = 1.33: 1) [v]/[v̥]/[w] (Araki, Mpotovoro);

EV = 1.5: 1) [v]/[w] (Hiw, Loh, Lehali, Lehalurup, Motlav, Vatrata, Mota, Mosina, Lakona, Nume, Koro, Wetamut, Merlav, Nokuku, Wusi, Marino, Sowa, Atchin, Maragus, Labo, Namakura, Lenakel, Whitesands, South-west Tanna, Kwamera), 2) [f]/[w] (Lonwolwol, South Efate), 3) [b]/zero (Sa);

EV = 2.0: 1) [v]/[w]/zero (Valpei, Tasmate, North Tanna), 2) [v]/zero (Raga, Seke), 3) [h] (South-east Ambrym);

EV = 2.33: 1) [h]/[w]/zero (Aneityum);

EV = 2.5: 1) [y]/zero (Sakao).

The mean erosion value for PAN \*p in the languages of Vanuatu is thus  $(9 \times 0.5) + (44 \times 1.0) + (2 \times 1.33) + (28 \times 1.5) + (6 \times 2.0) + (1 \times 2.33) + (1 \times 2.5) = 110/91 = 1.209$ .

#### NEW CALEDONIA AND THE LOYALTY ISLANDS

In some of the languages of New Caledonia PAN \*p is retained as a stop in initial position, but never in intervocalic position. Erosion values are:

EV = 1.0: 1) [v] (Voh-Kone, Iaai), 2) [f] (Faga Uvea), 3) [p]/[v]/[w] (Kumak);

EV = 1.5: 1) [p]/zero (Caac, Jawe, Nemi, Fwâi, Pije, Xârâcùù, Xârâgùre).

The mean erosion value for PAN \*p in the languages of New Caledonia and the Loyalty Islands is thus  $(4 \times 1.0) + (7 \times 1.5) = 14.5/11 = 1.318$ .

#### ROTUMA, FIJI AND TRIANGLE POLYNESIA

Although the nasal grade of PAN \*p is commonly reflected as a stop in the languages of this geographical region, the oral grade of \*p has invariably changed.<sup>6</sup> Erosion values are:

EV = 1.0: 1) [v] (Western Fijian, Eastern Fijian), 2) [f] (Tongan, Niue, Samoan, Tuvaluan, Wallisian, Tahitian, South Marquesan) 3) [ɸ] (Moriiori);

EV = 1.5: 1) [f]/[h] (Tuamotuan), 2) [ɸ]/[h] (Maori).

EV = 2.0: 1) [h] (Rotuman, Manihiki-Rakahanga, Tongareva, North Marquesan, Hawaiian), 2) [ʔ] (Rapa, Austral, Rarotonga, Mangarevan), 3) [w] (Pukapukan).

The mean erosion value for PAN \*p in the languages of Rotuma, Fiji and triangle Polynesia is thus  $(10 \times 1.0) + (2 \times 1.5) + (10 \times 2.0) = 33/22 = 1.500$ .

Table 2 summarises the results obtained in this survey of the data. (1) = mean erosion value, (2) = distance from Taiwan in kilometres. Mean erosion values are given both as a numerical expression and as an approximation to a phonetic quality. Distances were calculated by approximating the mid-point of each geographical region, and measuring from this to central Taiwan. Geographical categories in Table 2 are arranged by increasing distance from the probable Austronesian homeland.

TABLE 2: SPECIFIC EROSION VALUES FOR PAN \*p IN RELATION TO DISTANCE FROM TAIWAN

No.	AREA	(1)	(2)
1.	Taiwan	.133 (p)	-
2.	Philippines	.145 (p)	1230
3.	Mainland SEA	0 (p)	1930
4.	Borneo	.023 (p)	2810
5.	Sulawesi	.118 (p)	2900
6.	Moluccas	1.605 (f/h)	3160
7.	Micronesia	1.600 (f/h)	3800
8.	Sumatra-Sumbawa	.194 (p)	3865
9.	Lesser Sundas	1.194 (f)	3865

<sup>6</sup>For purposes of determining erosion values I have disregarded the Eastern Polynesian 'double labial dissimilation' (as in PPN \**faiŋe* > PEPN \**wahine*). However, since [w] and [h] have the same numerical value in my interrelation of the erosion scale of \*p, this simplification of the data has no effect on the outcome of the calculations.

10.	New Guinea	1.348 (f)	4390
11.	Bismarck Archipelago	.925 (f)	4680
12.	Solomons-Santa Cruz	1.122 (f)	5560
13.	Vanuatu	1.209 (f)	6730
14.	New Caledonia	1.318 (f)	6900
15.	Madagascar	1.000 (f)	8243
16.	Rotuma, Fiji, Polynesia	1.500 (f/h)	10200

What, if anything, does the summary in Table 2 tell us about a possible correlation between erosion value and migration distance? The regions have been arranged by increasing distance from Taiwan, and in 10 of the 15 transitions between adjacent categories the erosion value increases. This in itself suggests a positive correlation between degree of phonological erosion and mean lineal distance from Taiwan. However, in a number of cases (Taiwan-Philippines, mainland Southeast Asia-Borneo, Borneo-Sulawesi, Solomons-Vanuatu, Vanuatu-New Caledonia and Loyalties) these increases in erosion values are statistically insignificant.

If instead we group the mean erosion values of Table 2 into larger geographical categories defined by thousand kilometre intervals we see a much more focused pattern (Table 3).

TABLE 3: EROSION VALUES OF \**p* AS A FUNCTION OF DISTANCE FROM TAIWAN IN 1000KM INTERVALS

DISTANCE	MEAN EROSION VALUE
zero-1000	-
1000-2000	.073
2000-3000	.071
3000-4000	1.148
4000-5000	1.115
5000-6000	1.122
6000-7000	1.264
7000-8000	-
8000-9000	1.000
9000-10000	-
10000-11000	1.500

In Table 3 an abrupt leap in erosion value appears at about 3000km distance from Taiwan. Languages less than this distance from the probable Austronesian homeland overwhelmingly reflect PAN \**p* as [p], whereas the norm for those located more than 3000km from Taiwan is [f], with a value mid-way between [f] and [h] in Polynesia. There are, however, serious difficulties with reaching straightforward conclusions from these figures.

First, lineal distance from a probable homeland and actual migration distance are not the same thing. Many of the Oceanic languages of Micronesia are no more than 3800km from Taiwan, while those of New Guinea and the Bismarck Archipelago are somewhat more distant. Yet languages such as Trukese, Ponapean or Marshallese are descended from Proto Oceanic, which probably was spoken in the New Guinea-Bismarck Archipelago region. In this case, then, greater migration distance has *decreased* lineal distance from the probable Austronesian homeland. In principle we could correct for this discrepancy between lineal distance and migration distance if we knew the migration routes

which led to the settlement of each area represented in Table 2. Unfortunately, however, our knowledge of these matters is at best highly tentative and probabilistic.

A second difficulty with accepting the correlation of mean erosion value and lineal distance in Table 3 is that there is essentially no gradient within either of the sharply distinguished geographical regions. In other words, there is no evidence that erosion value increases with lineal distance for the languages of Taiwan, the Philippines and western Indonesia. Similarly, apart from Polynesia, there is little evidence that erosion values increase with increasing distance from the Moluccas to New Caledonia and the Loyalties. What can hardly be overlooked in this pattern, however, is the abrupt increase of erosion values when we pass from languages of the Western Malayo-Polynesian group to languages of the Central-Eastern Malayo-Polynesian group. Apart from Madagascar, which must be discarded as statistically unusable on grounds of sample size, no WMP-speaking area has an erosion value for \**p* that even remotely approaches 1.0, whereas erosion values above 1.0 are the norm throughout the 580-language CEMP-speaking area of eastern Indonesia and the Pacific. This apparent correlation of erosion value with major subgroup boundary as opposed to lineal distance is starkly evident in comparing geographical regions 8 (WMP-speaking) and 9 (CEMP-speaking), which are about equally distant from Taiwan.

Why does the erosion value of \**p* appear to correlate with major subgroup boundaries? The facile answer to this question would be that \**p* had already begun to lenite in Proto Central-Eastern Malayo-Polynesian. Yet many CEMP languages retain PAN \**p* as a voiceless bilabial stop. Are all of these stop reflexes instances of secondary fortition? Or have there been factors other than migration distance which have contributed to the higher erosion values of \**p* in CEMP-speaking areas? And if so, what might these be?

In his contribution to this volume Ross has found a positive correlation between what he calls 'sedentariness' and 'conservatism'. He has discovered this correlation from a compilation of *all* sound changes in all members of a small group of closely related languages. Although the correlation he discusses appears to be well-founded, Ross is well aware that there is no obvious reason why it should exist. In an effort to explain it he proposes a probabilistic scenario of how language communities might hive off to form new linguistic colonies. According to Ross, emigrants from an established language community are likely to consist predominantly of members of the younger generation together with their near relatives (young married couples and their children), while those left behind would be mostly elderly:

If Oceanic settlement occurred through small groups who left their home villages on the periphery of the region in which Oceanic was spoken and moved to a place beyond that periphery, then it is probable that the settlers were usually led not by the gerontocrats of the village, but by the middle generation of men with young families. The new settlement would then have a different generational and authority structure from the old. As a result the speech patterns of the younger generation, which tend in most cultures to be innovative, would form a larger proportion of daily discourse than before, and, in the absence of the former elders, would go unchecked. Change resulting from this situation would accord with the probably innovative values of the new migrants, and, as the new community developed its own sense of identity, more novel forms of speech would become emblematic of communal identity and thereby entrenched. Since in many cases the new community would have maintained contact with the old, members of the old community would be aware of this linguistic innovation, and would emphasise the conservative features of their own speech as emblems of *their* identity.



What is noteworthy about the macrocomparison conducted here is that it agrees with Ross in revealing a general correlation between amount of sound change and lineal distance from the probable homeland, but the explanation for this correlation cannot be the one which Ross proposes. PAN *\*p* must have remained a voiceless bilabial stop for many generations after the break-up of Proto Central-Eastern Malayo-Polynesian, yet CEMP languages exhibit far higher erosion values of *\*p* than do Formosan or WMP languages. Given this observation a correlation of sound change with migration distance can be maintained only by arguing that every CMP language arrived in eastern Indonesia and that every OC language arrived in the Pacific as a result of separate migrations – one per attested language. If we reject this scenario (which clearly we must) it follows that the far higher erosion values of *\*p* in CEMP languages are the result of changes which in some cases must have taken place *many generations after* the major moves which brought the languages to or near their historic locations. If indeed there is a causal relationship between migration and amount of innovation we would expect change in a language to follow immediately upon the physical relocation of its speakers. On the level of microcomparison explored by Ross in his paper this appears to be the case; on the level of macrocomparison explored in this paper it does not.

#### 4. CONCLUSION

In conclusion, mapping of the erosion sequence *\*p* > [f] > [h] > zero in Austronesian languages as a whole reveals a general correlation between erosion value and lineal distance from the probable Austronesian homeland. However, this correlation appears to be largely a by-product of factors other than distance as such. Ross's proposal that younger and linguistically more innovative speakers would tend to venture to new homelands, leaving the older and linguistically more conservative generation behind, is difficult to apply to broad patterns of phonological erosion, since many generations must have separated speakers of Proto Malayo-Polynesian (in which PAN *\*p* was reflected as *\*p*) and Proto Oceanic (in which PAN *\*p* was also reflected as *\*p*). Unless PAN *\*p* had already begun to lenite in PCEMP, and underwent secondary fortition far more often than is generally assumed, we are confronted with a correlation that has no ready explanation: *\*p* exhibits a drift-like tendency to erode at a far more rapid rate in CEMP languages than in WMP languages.

Before concluding we might note impressionistically that several other erosion sequences in Austronesian languages appear to follow a mapping pattern very similar to that of *\*p*. Among these are: 1) *\*S* > [s] > [h] > zero (sibilant reflexes only in Taiwan, [h] fairly common in the Philippines and less so in western Indonesia, with non-zero reflexes (*h-*) in a single CEMP language; cf. Blust 1981), 2) *\*k* > [h], [ɣ], [ʔ] > zero (eroded reflexes virtually absent in Taiwan, rare in the Philippines and western Indonesia, but relatively common in eastern Indonesia and the Pacific), 3) *\*t* > [s], especially before *\*i* (absent in Taiwan, rare in the Philippines and western Indonesia, common in the Pacific), and 4) loss of final consonants (absent in Taiwan, rare in western Indonesia, not uncommon in eastern Indonesia, extremely widespread in the Pacific).

The geographical distribution of these innovations still must be carefully documented. However, once the documentation is available it appears likely that the question we have been forced to ask in relation to *\*p* will return even more insistently to demand an answer: is the tendency to increased phonological change in the Austronesian languages of eastern Indonesia and the Pacific as compared with those of Taiwan, the Philippines and western Indonesia purely a function of increased migration distance, or is the correlation of this spatial dimension with the WMP:CEMP language boundary indicative of some other causal factor or factors that continue to elude us?



## BIBLIOGRAPHY

- BENDER, Byron W., Robert W. HSU, Frederick H. JACKSON, Jeffrey C. MARCK, Kenneth L. REHG, Ho-min SOHN, Stephen TRUSSEL and Judith W. WANG**, 1984, Micronesian cognate sets. Computer printout. Honolulu, Department of Linguistics, University of Hawaii.
- BENEDICT, Paul K.**, 1984, Austro-Tai parallel: a tonal Chamic language on Hainan. *Computational Analyses of Asian & African Languages*, 22:83-86.
- BERG, René van den**, 1990, Muna dialects and Munic languages: towards a reconstruction. In Ray Harlow and Ross Clark, eds *VICAL 2: Papers from the Fifth International Conference on Austronesian Linguistics*. Auckland: Linguistic Society of New Zealand.
- 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: Eastern Austronesian*. PL, C-61:691-716.
- BLUST, Robert**, 1976, Review of Dahl, *Proto-Austronesian*. *Language* 52/1:221-237.  
1981, The Soboyo reflexes of Proto-Austronesian \*S. In Robert A. Blust, ed. *Historical linguistics in Indonesia*, Part I, 21-30. NUSA: Miscellaneous Studies in Indonesian and Languages in Indonesia, vol.10. Jakarta: NUSA.  
1984/85, The Austronesian homeland: a linguistic perspective. *Asian Perspectives* 26:45-67.  
n.d., Fieldnotes on languages of Borneo, the Lesser Sundas and western Melanesia.
- COLLINS, James T.**, 1982, 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.
- CONANT, Carlos Everett**, 1908, "F" and "V" in Philippine languages. *Bureau of Science, Division of Ethnology Publications* 5/2:135-141.
- DAHL, Otto Chr.**, 1951, *Malgache et Maanjan*. Oslo: Egede-Instituttet.  
1976, *Proto-Austronesian*. 2nd edn. Scandinavian Institute of Asian Studies monograph series, No.15. London: Curzon Press.
- FEINBERG, Richard**, 1977, *The Anutan language reconsidered: lexicon and grammar of a Polynesian Outlier*. 2 vols. HRAFlex Books, OT1-001. New Haven, Conn., Human Relations Area Files, Inc.
- FERRELL, Raleigh**, 1969, *Taiwan aboriginal groups: problems in cultural and linguistic classification*. Institute of Ethnology, Academia Sinica, monograph No.17. Taipei.
- FOX, James J.**, n.d., Lexicostatistical lists for languages of the Lesser Sunda islands.
- HAUDRICOURT, André G.**, 1971, New Caledonia and the Loyalty Islands. In Thomas A. Sebeok, ed. *Current trends in linguistics*, vol.8: *Linguistics in Oceania*, 359-396. The Hague: Mouton.
- HAUDRICOURT, André G., and Françoise OZANNE-RIVIERRE**, 1982, *Dictionnaire thématique des langues de la région de Hienghène (Nouvelle-Calédonie)*. LACITO- documents, Asie-Austronésie, 4. Paris.
- HIMMELMANN, Nikolaus P.**, ed., 1990, Sourcebook on Tomini-Tolitoli languages. MS.
- HUDSON, Alfred B.**, 1967, *The Barito isolects of Borneo*. Southeast Asia Program, Data Paper No.68. Ithaca, New York: Cornell University Press.
- LEE, Ernest Wilson**, 1966, Proto-Chamic phonologic word and vocabulary. PhD dissertation, Indiana University. Ann Arbor, University Microfilms International.
- LEES, Robert B.**, 1953, The basis of glottochronology. *Language* 29/2:113-127.
- MARSH, Mikell Alan**, 1977, Favorlang-Pazeh-Saisiat: a putative Formosan subgroup. PhD dissertation, Washington State University. Ann Arbor, University Microfilms International.
- MATISOFF, James A.**, 1990, On megalocomparison. *Language* 66/1:106-120.
- McFARLAND, Curtis D.**, 1977, *Northern Philippine linguistic geography*. Studies of Languages and Cultures of Asia and Africa Monograph Series, No.9. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

- MILLS, Roger F., 1975, Proto South Sulawesi and Proto Austronesian phonology. 2 vols. PhD dissertation, University of Michigan. Ann Arbor, University Microfilms International.
- PRENTICE, D.J., n.d., Comparative vocabularies for languages of Sabah.
- RAY, Sidney H., 1913, The languages of Borneo. *Sarawak Museum Journal* 1/4:1-196.
- REID, Lawrence A., ed., 1971, *Philippine minor languages: word-lists and phonologies*. *Oceanic Linguistics* Special Publication No.8. Honolulu: University of Hawaii Press.
- ROSS, Malcolm D., 1988, *Proto Oceanic and the Austronesian languages of western Melanesia*. PL, C-98.  
in this volume, How conservative are sedentary languages? Evidence from western Melanesia.
- SNEDDON, J.N., 1978, *Proto-Minahasan: phonology, morphology and wordlist*. PL, B-54.  
1984, *Proto-Sangiric and the Sangiric languages*. PL, B-91.
- STRESEMANN, Erwin, 1927, *Die Lauterscheinungen in den ambonischen Sprachen*. *Zeitschrift für Eingeborenen-Sprachen*, Supplement 10. Berlin: Reimer.
- 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.
- TSUCHIDA, Shigeru, 1971, List of words of Formosan languages. MS.  
1982, *A comparative vocabulary of Austronesian languages of Sinicized ethnic groups in Taiwan*, Part I: West Taiwan. Memoirs of the Faculty of Letters, University of Tokyo, No.7. Tokyo.  
1985, Kulon: yet another Austronesian language in Taiwan? *Bulletin of the Institute of Ethnology, Academia Sinica* 60:1-59.
- VÉRIN, Pierre, Conrad P. KOTTAK and Peter GORLIN, 1969, The glottochronology of Malagasy speech communities. *Oceanic Linguistics* 8/1:26-83.
- WURM, S.A. and Shirô HATTORI, eds, 1981, *Language atlas of the Pacific area*, Part I: New Guinea area, Oceania, Australia. Canberra: Australian Academy of the Humanities in collaboration with the Japan Academy. PL, C-66.
- ZORC, R. David, 1974, Internal and external relationships of the Mangyan languages. Papers of the First International Conference on Comparative Austronesian Linguistics. *Oceanic Linguistics* 13:561-600.

## PROTO OCEANIC CULTURE: THE EVIDENCE FROM MELANESIA

ANN CHOWNING

### 1. INTRODUCTION

It seems safe to say that the vast majority of linguists interested in Austronesian (AN) languages now accept the existence of a division of them usually called Oceanic, though there still remains some uncertainty about all the details of its borders.<sup>1</sup> Nevertheless, apart from a few languages of western Irian Jaya (Grace 1976:62; Pawley 1981:300), it is agreed that the AN languages of Melanesia are all OC. Here I am defining Melanesia as extending from New Guinea and the Bismarck Archipelago to Fiji and including all the intervening islands except those few on which Polynesian languages are spoken (see Chowning 1977). Admittedly the AN status of a number of Melanesian languages is disputed, and one of these is a language with which I have worked, Sengseng (Chowning 1985). For most languages, however, the classification as AN or NAN is clearcut.

It has not, however, been so easy to correlate culture traits with the linguistic distinctions. One reason is that Melanesia does not constitute a single unified culture area like Polynesia or Micronesia. Nevertheless, a number of culture traits are widespread, though not universal, in Melanesia, and attenuated, rare or non-existent in Polynesia and Micronesia, while a number of characteristic Polynesian and Micronesian traits are rare or absent in Melanesia (see Chowning 1977; Pawley 1981:298-299). These characteristic Melanesian traits cannot, however, easily be correlated with speakers of OC languages within the boundaries of Melanesia, because most of the traits are also found in the cultures of speakers of NAN languages who occupy most of western Melanesia. In most cases it is no longer possible to attribute the origin of some of these shared traits to the ancestors of the present OC speakers. If one does not accept Terrell's (1986) radical suggestions about the artificiality of the distinction between AN and NAN languages, with its implications for the time that AN speakers appeared in Melanesia, then the cultivation of root crops such as taro and the manufacture of pottery probably existed in western Melanesia well before AN speakers arrived there (see Bulmer 1982). What can be said, however, and with reasonable certainty, is that the earliest AN speakers in Melanesia possessed these traits along with various others. The evidence is of two sorts:

---

<sup>1</sup> Abbreviations used are as follows: EO – Eastern Oceanic, MN – Melanesian, NAN – non-Austronesian, OC – Oceanic, PAN – Proto Austronesian, PEMP – Proto Eastern Malayo-Polynesian, PEO – Proto Eastern Oceanic, PMN – Proto Melanesian, PMP – Proto Malayo-Polynesian, PNGA – Proto New Guinea Austronesian, PNH – Proto North Hebridean, PPH – Proto Philippines, PPN – Proto Polynesian, PWMN – Proto Western Melanesian, PWMP – Proto Western Malayo-Polynesian, PWO – Proto Western Oceanic, WMP – Western Malayo-Polynesian, WOC – Western Oceanic.

---

Robert Blust, ed. *Currents in Pacific Linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, 43-75. *Pacific Linguistics*, C-117, 1991.

the distribution of the traits throughout Melanesia and, often, well beyond, and comparative linguistics. It is the combination of these that make clear the association between AN speakers and both the outrigger canoe and distinctions of relative age for same-sex siblings. These are not, however, peculiarly Melanesian traits, and here I am particularly concerned with those that may characterise the Melanesian region, or parts of it.

With acceptance of the validity of OC came attempts to reconstruct a specifically POC lexicon. A brief history of most of these attempts is given in Pawley and Green (1984). An extremely useful early compilation was Grace's *Proto-Oceanic finder list* (1969, following a cyclostyled version of 1965), which included data proposed by Biggs, Milke, Goodenough, Capell and Grace himself. Grace's continuing interest in this field is shown in the computer list he and his associates have assembled at the University of Hawaii. (I have seen only an 'edited' 1979 printout of this, variously attributed to Grace himself and to Peter Lincoln, and it is this version to which I shall be referring below, as Grace-Lincoln.) Over the years, many linguists have added POC reconstructions, but a considerable number of these are scattered in articles or, worse, in personal files, and at present it is impossible to say what the total reconstructed POC lexicon is.<sup>2</sup> Furthermore, it is also difficult to be sure how many of the forms assigned to POC actually belong to some smaller subgroup. When, for example, Pawley and Green (1984:128) say that a "reconstruction must be represented by cognates in at least two first-order subgroups of Oceanic", the problem is that ideas about the composition of first-order subgroups change with considerable frequency, and vary from one linguist to another. It must consequently be understood that in some cases, lexemes assigned to POC may actually belong only to a lower level proto-language. This fact was clearly recognised in Grace (1969), where forms originally ascribed only to PEO, PNGA, and PPN were signalled. Grace explained: "The reasoning is that such reconstructions, where sound, represent forms of some antiquity, and that it is therefore a possibility that they were in the PO [Proto Oceanic] vocabulary" (1969:40). Over time, however, it has become commonplace for the label POC to be used without discrimination for any terms found in Oceania and not wholly confined to Polynesia (or Micronesia). Publication of proto-forms often does produce examples from a wide scatter of languages and so supports the assignment of any particular lexeme, once apparently confined to a limited region, to POC itself. It is nevertheless dangerous to attribute to 'POC speakers' traits for which linguistic evidence is limited to a portion of the OC-speaking region<sup>3</sup> (Pawley and Green 1984:130).

Contrary to what I once suggested (Chowning 1963), I am not assuming that the OC languages of Melanesia form a closed genetic group. Like Pawley (1981), however, I find it useful for the reconstruction of a portion of culture history to deal with the OC languages of Melanesia as a separate subset of OC, and I shall use Proto Melanesian in that sense. At the same time, I shall not assume that cognate forms found only in Melanesia should automatically be ascribed to POC, or to PMN. Some of them may be local developments – widespread, but not known to occur throughout the region. Terms for slit gong are a case in point (see below). At best, it may be that some forms that have been or could be labelled POC more properly belong only to a western region, or only to an eastern one. Here I shall use 'western Melanesia' to designate the region from western New Guinea

<sup>2</sup>Certainly some duplication of effort is involved. This paper was almost finished before I discovered that Blust too, admittedly in 1984, had reconstructed words for 'doorway' and 'taro pudding' which I thought I was the only one to have noticed.

<sup>3</sup>In some cases, data that suggest a wide distribution are dubious. Grace-Lincoln cite Bola *kaw* in reconstructing a POC word for kava, together with a Gedaged form, *aiu*, which looks much more dubious. I assume that these two constitute the grounds for assigning the proto-form to POC (see Pawley and Green 1984), but to the best of my knowledge Bola words never end in a consonant (nor is kava used on the Willaumez Peninsula).

through the western Solomons, including and ending at the boundaries of Ross's (1988) North-West Solomonian group, so that it encompasses the equivalent of his 'Western Oceanic', but also, temporarily at least, includes the Admiralties and Mussau-Emira, and 'eastern Melanesia' for the rest of the region. The latter should not be understood to imply that all, or almost all, the eastern languages belong together in a group that can be labelled Eastern Oceanic.

Relatively few of the linguists who have been reconstructing POC have also been interested in culture history rather than simply in linguistic relationships. Even for the best-studied part of Oceania, Polynesia, Clark (1979:268) points out that "[b]oth material culture and social organization in the Proto-Polynesian lexicon have received very little linguistic attention", and that even kinship terminology has not been fully studied. Interestingly, kinship is the one area that has received a great deal of attention in Melanesia, even if disagreements and questions remain (see below). Otherwise not a great deal has been done to use linguistic evidence to reconstruct the culture of the early speakers of OC languages in Melanesia. In their recent survey of POC reconstructions, Pawley and Green (1984:131) point out that the "list of artifacts ... is not very impressive." In fact, their list is unnecessarily short; for example, it contains 'needle' but not 'sew' (*\*saqit* and perhaps other terms), and 'clay pot' but not 'stone oven' (*\*qumun*). Many forms found in both Melanesia and Polynesia, or in Melanesia alone, have been reconstructed but not brought together to give insights into what might have been the culture shared by speakers of POC before the development of PPN. This is a preliminary attempt to remedy that situation.

The procedure followed here will be conservative. If a term is found both in Indonesia and in Polynesia or Micronesia, I shall assume that it once existed in Melanesia even if it no longer seems to. Otherwise, I shall call a term PMN only if it is attested from parts of Melanesia geographically distant from each other, preferably in both eastern and western Melanesia, and not to date assigned to the same subgroup of languages, or if it is found in Melanesia east of the Sarmi Coast and in some other part of the AN-speaking region.<sup>4</sup> If all of the forms known to me should come, say, only from the region of New Guinea and the Bismarck Archipelago, I shall for the present call the proto-form only, western Melanesian; if only from the region from the southern Solomons eastwards, eastern Melanesian. Smaller regions will also be noted in a few cases.

Furthermore, since this is an exercise in cultural reconstruction, I shall be careful about deducing cultural practices from terms that have a range of meanings (see section 2.1.3). Also, as Pawley and Green (1984:131) point out, there is archaeological evidence in various cases for traits for which linguistic evidence is lacking. If we make the usual assumption that the bearers of Lapita culture spoke early forms of OC, if not POC itself, we might expect to find POC words for 'obsidian' and 'shell bracelet', but at present these are not reconstructible (see discussion below of the term for trochus shell). We also know because of the distribution of certain traits in later Melanesian societies that POC words for which there is no evidence must have existed. A conspicuous example is the absence of a word for 'betel pepper' when it is possible to reconstruct other terms that relate to chewing betel.

The reasons that this is only a first attempt are several. First, I have not yet had time to engage in a systematic search and comparison of, for example, words for 'hunt' and 'club' in different languages. This means that my data are biased either towards checks of the distribution of forms that

<sup>4</sup>For example, I would attribute *\*misa*/*\*ma-isa* 'only, alone' to POC solely because Sengseng *misa*- (which takes a suffixed possessive) agrees in shape and meaning with Blust's (1986) reconstruction of *\*ma-isa*, with no Oceanic witnesses, for PMP.



have been reconstructed by others, or towards forms that resulted from my recognising in reading about other languages possible cognates with words in the languages that I speak (Lakalai, Sengse, Kove of West New Britain, and Molima of Milne Bay, all in Papua New Guinea). Another reason is the dearth of published comparative material, and the fact that it has been difficult for me to obtain some of what does exist. With the rise of interest in lexicostatistics has come a dependence on 'non-cultural' vocabulary, and many of the published lists (such as those in Tryon and Hackman 1983) contain almost none of the lexemes that interest me. Some of the much older material, from Codrington and Ray to Churchill, is occasionally more useful, though greatly restricted by the dearth of sources available to them and by, as regards Codrington and Ray, their focus on forms with apparent cognates in Indonesian languages. The only complete published list dealing with cultural items is Leenhardt's for New Caledonia (1946), but most of these forms seem to be peculiar to New Caledonia.

The consequence is that I have relied heavily on a limited set of dictionaries and word lists – the published ones, my own manuscripts for the languages mentioned above, and a few manuscripts made available to me by colleagues, particularly Andy Pawley. In addition, I have been grateful for copies of unpublished data from Malcolm Ross and Bob Blust. I have found fragments of material in other sources such as ethnographies, but realise that there is much more that I have not seen, or recognised. It is hoped that others will be able to expand on, and emend, the proto-forms suggested here. Throughout, I shall use *G* to represent the velar fricative.

## 2. CULTURAL ITEMS

### 2.1 MATERIAL CULTURE

#### 2.1.1 CLOTHING, ORNAMENTS AND OPERATIONS ON THE BODY

These are areas of enormous variation from one society to another. Although ornaments were worn everywhere, clothes were not; for example, in parts of the Bismarck Archipelago both sexes wore nothing (Parkinson 1907). In others, as Lakalai and part of Kove, only women wore clothing. Where it did exist, it was usually quite different for men and women, hence the impossibility of reconstructing a POC term for clothing in general. Nevertheless, two terms have been reconstructed, both assignable to POC. One is *\*malo*, probably 'male garment of barkcloth'. The shape of the garment ranges from a strip passed between the legs to cover the genitals, or sometimes the penis alone (the so-called 'G-string'), to a wrap-around kilt, but except for the use of the same term to cover 'barkcloth' and the trees from which it is made, the range of meanings is restricted. Barkcloth used for purposes other than clothing is often called by a different term (e.g. Lakalai *malaha* 'barkcloth sling for carrying a baby', *voi* 'barkcloth for mask'). Very rarely were female garments made of barkcloth; more often, the basic material was shredded leaves (rarely if ever grass, despite the usual misnomer 'grass skirt'). I have not been able to reconstruct a term for a female garment, though for the west there are suggestive resemblances between Motu *rami* (with cognates in other Central Papuan languages) and Gitua *rami*. The other POC term, *\*tipi*, is glossed 'belt' in Grace-Lincoln, but in Central Papua, where reflexes of *\*malo* have not been recorded, it typically designates the male garment, and is so interpreted in Capell (1943). Its range of meaning, however, includes 'female garment', 'baby sling' and 'native cloth'. I suspect that it may originally have been a generic term for 'barkcloth' – perhaps, in view of its shape, a doublet of PPN *\*tapa* (or related to *\*tapis* 'cloth', reconstructed by Blust?). I have not been able to reconstruct, for Melanesia, any of the terms relating to the manufacture of barkcloth. The verb for beating it out usually reflects either POC



\**katu*, as in Lakalai, or POC \**tutua*/\**tutuki*, as in Sengseng. (Grace-Lincoln give cognates derived from the latter only for eastern Oceania and PAN, but Lakalai, Kove and Molima all contain *tutu*, and Sengseng *tut*, with meanings ranging from 'tap' to 'beat'.)

Perhaps the most common everyday ornaments were necklaces of shell or animal teeth, bracelets and armbands of a variety of materials, most often worn above the elbow, and ear ornaments. Noses were usually pierced, but it is not clear how often nose ornaments were worn. The other very common ornaments were legbands, worn below the knee, and anklets; belts and girdles; and hair ornaments, including ornamental combs. Aromatic herbs were commonly stuck onto armbands and into the hair; feathers too, unmodified or made into simple ornaments, were often stuck into the hair. Obviously it is not possible to reconstruct terms for the simplest ornaments, or to distinguish between, say, the usual word for 'comb' and one indicating its use as an ornament. Nevertheless, it is surprising that it has been possible to reconstruct so few terms. Where bracelets were made of shell or tortoiseshell, the words for both the artefact and the material were often the same. Consequently, although \**lala* 'trochus shell' has been reconstructed for POC, and the cognate terms often designate bracelets of this material, it is difficult to know whether the 'bracelet' meaning can be attributed to the proto-form, and indeed Grace-Lincoln have not done so, despite the archaeological evidence mentioned above.

One of the most distinctive Melanesian ornaments is the so-called *kapkap*, from its name in a New Ireland language (Reichard 1933), composed of tortoiseshell cut into intricate patterns and fastened to a disk of white shell. Despite its wide distribution, at least from the Admiralties through the southern Solomons (and see the similar ornaments of the Marquesas), no set of cognate names for it has been reconstructed.

In the key to his 1969 list, Grace gives \**sa(d,r)u* 'head ornament', with another meaning 'comb', taken from Milke. In fact the evidence for this head ornament gloss does not strike me as strong, with only two cases, Motu *daria* 'feather ornaments of comb', Gedaged *saz* 'headgear of feathers'. Beside Sia *saraka*, Sariba *suari* 'comb', there exists Manam and Sepa *saru* 'comb'. In reconstructing the POC word, I suspect that Milke was originally influenced by his knowledge of the Fijian and Polynesian words for 'comb', which he cites (Fiji *seru*, PPN \**selu*). For the present, the forms justifying \**sa(d,r)u* belong to a very limited area, geographically, and cannot be called even western Melanesian. Blust (1980), however, reconstructs PMP \**saRu* (and other words for 'comb'), and cites Samoan *au*, so presumably this form should be preferred to \**sa(d,r)u*. In any case, the 'head ornament' gloss should be dropped. The comparative evidence for POC reflexes of PMP \**suat* 'comb' are also good; Ross (n.d.) has reconstructed Proto Papuan Tip \*(*n*)*suad* to account for the many cognates in that region of PNG.

Another proto-form that probably related to bodily decoration is \**pani* 'to apply oil or paint to the head or body'. The range of meaning varies from language to language; in some cases reflexes seem to refer only to the painting of objects, as with Sengseng *pan*, Samoan *vali*, both of which are used for painting designs on barkcloth. But with examples such as Lakalai *vali*, 'to apply paint, feathers, or other surface adornment to the head or body' and Fiji *wali* 'to anoint, to oil hair', I suggest that a prime meaning of PPN \**pani* 'to paint, smear, rub oil on' had to do with bodily adornment. Clearly the term itself, however defined, can be ascribed to PMN, and indeed to POC.

Terms for 'shave' and 'cut hair' tend to have so many wider, though related, meanings that it is difficult to suggest that one of these meanings was primary. Unfortunately, the same problem applies to the terms for most operations on the body. The most common were probably piercing of the nasal

septum and slitting (rather than piercing) of the earlobes, so that ornaments could be inserted. The verbs, however, tend to be general terms for 'pierce' and 'slit/cut'. Penile operations were rarer, and of two sorts: circumcision, and super(in)cision, in which the foreskin is simply slit and retracts of its own accord as the wound heals. Specialised terms exist for these operations in many areas, but the terms recorded are not usually cognate with each other or with PPN *\*tefe* 'circumcise' (usually in fact supercision), *\*kula* 'circumcision'. The exception is in some languages of Vanuatu; cf. Efate *tefi* 'to circumcise', with cognates from other languages cited by Churchill (1911:265) such as Malekula *teve*. In some languages the word is still a general one for 'cut', and in some (such as Mota) only means 'cut' (with a derivative meaning 'knife'). It still seems justifiable to attribute *\*tepe* 'to circumcise' to the Central Pacific, but not yet to a wider region.

Tattooing has a limited distribution in Melanesia, tending to be confined to people with relatively light skin, though not all of these tattooed. (It was, for example, rare in the Massim region of eastern New Guinea.) Where it was practised, methods ranged from the cutting of fine lines with obsidian, after which pigment was rubbed in, to the use of special implements. Again, it has not been possible to reconstruct a PMN term. Darker-skinned people practised ornamental cicatrisation instead of tattooing, but the terms for it tend to be those for 'cut' (the burning of ornamental scars was also common but not a major form of ornamentation). I have not made a study of the literature on tattooing techniques, but as far as I know the implements used in Melanesia did not include the typically Polynesian toothed chisel (see Buck 1958:296). For example, in Central Papua the skin was pierced by thorns set into a handle (Seligman 1910:265). The difference is of some importance because Green attributes tattooing to the whole Lapita area on the basis of "tattooing chisels from early Eastern Lapita contexts on Tongatapu" (1979a:16). He points to the reconstruction of PPN terms not only for the process (*\*tatau*) but for a kind of tattooing chisel (*\*hau*), and has reconstructed a special word, *\*uhi*, for the "Eastern Polynesian tattooing chisel". (I do not know if there are differences in the implements or only in the words.) Unless the terms used for implements associated with cicatrisation in regions like the Solomons are cognate with some of these terms, it looks for the present as if we lack linguistic confirmation of tattooing as an early trait throughout Melanesia. It should also be noted that Blust considers PPN *\*hau* to be derived from POC *\*saRu* 'comb'. Interestingly, a somewhat archaic Kove term for tattooing, *totoriŋa*, appears to be cognate with PPN *\*tosi* 'mark, draw line'. Geraghty (1983:141) defines this as 'score a line' and notes cognates in Fijian and Arosi. Nevertheless, without confirmation from other parts of Melanesia I can not assume that 'tattoo' was a general meaning for a term related to *\*tosi*, though I am willing on this evidence to call the term POC.

One other bodily operation was widespread: the blackening of teeth. Early observers often confused deliberate blackening with the accidental effects of betel-chewing, but lexicons as well as the accounts of closer observers make clear how common the practice was. The blackening material differed from area to area, being sometimes mineral and sometimes vegetable. Sometimes the word for it seemed to be related to the term for 'black', as in much of New Britain, where the mineral (probably manganese) was called something like *kit* or *keto*, a term possibly related to PAN *\*qitem* but possibly only to MN words such as Lakalai *kato* (and see also Nggela *katoa* 'very dark or dirty,

soot' as well as PPH *\*(Øq)ateq*.<sup>5</sup> There is, however, a word occasionally found in western Melanesia that refers only to tooth blackening – the process or the material: Tolai *taval*; Molima *tavana*; Roviana *davala*. These point to at least a proto-western term something like *\*ntapala*. Words used elsewhere, such as Proto Malaita *\*oko*, are unrelated. In view of the frequency of tooth-blackening in Indonesia, it is not a surprising trait to attribute to PMN culture. The materials were vegetal in the three regions for which the term derives from *\*ntapala*.

My data, incidentally, endorse the attributing of mirrors to PMN, even if they always involved only water as the reflective surface (in Lakalai, backed by charcoal in a stone basin). The word *\*tido* is glossed in Grace-Lincoln as 'to look at, especially in mirror' (derived from PAN). They cite only Fijian and Nggela within Melanesia, the latter glossed as 'to gaze' but Lakalai *titiro* 'mirror' confirms the distribution of this meaning throughout the region.

### 2.1.2 HOUSING AND RESIDENCE

I am able to add little to the usual words that have been reconstructed for houses and parts of them. The word, *\*katama* 'doorway', which I had reconstructed without knowing that Blust had also done so,<sup>6</sup> obviously adds nothing to our knowledge of PMN culture, since all houses must have doorways, and shelter is a human universal. Reflexes of the word are found all over Melanesia and also in Micronesia. Examples include Kove *atama*, Jabem *katam*, Rubi (of Central Papua) *atama*, *katama* in languages of Aoba (Ngwatua and Lolsiwoi), and Sonsorol *xatama*, with cognates in other Trukic languages (Bender n.d.). I assume that Molima *'atamana* and Efate *katema*, *ekatema*, which both mean 'outside' are derived from the same form, and also that the meaning 'doorway' is prior, both on distributional and logical grounds. Nevertheless, it must be noted that there is a widespread competing form for doorway, of which Lakalai *la mata la luma* is an example. In addition to most common meaning of 'eye', reflexes of *\*mata* very often refer to an opening, especially one through which light comes, such as the mesh of a net (see below). Presumably these meanings are related. In any case, it is common for a doorway to be called by a *mata* term.

Of the words that have been reconstructed for 'house', the standard two, *\*pale* and *\*Rumaq*, are both represented throughout Melanesia. I have not, however, found any evidence to support the suggestion in Blust (1984c) that *\*pale* specifically referred to public buildings. Sometimes, as with Tolai *pal*, the reflex is a generic term for all sorts of structures, including family houses, but in my data it is not the term for men's house or clubhouse, widespread though these are in Melanesia. I was unable to reconstruct a POC term for 'men's house', but Blust (pers. comm.) tells me that he has done so. The PMP term, *\*kamaliR*, has reflexes in the Admiralties as well as in the well-known PNH (*\*na*)*kamal*.

There is abundant evidence in Melanesia to support the POC terms for various parts of the house, such as 'sago thatch/roof' (*\*qatop*), 'wall' (*\*ndindi*) and 'rafter' (*\*kaso*); there is no need to support or challenge these reconstructions. My Lakalai material also supports Blust's (1980) reconstruction of PMP *\*kapit* 'fasten thatch with slats'. The case of 'ladder' is much more uncertain. Often houses

<sup>5</sup>Lakalai has reflexes of both, with *maheto* 'ripe *Canarium* almond' (with dark purple skin) and *kato* 'to paint black (for mourning), black mole on skin etc.'. It also has *keto* 'black stone found in the mountains'. Unlike many of their neighbours, including the Bola, the Lakalai did not blacken their teeth.

<sup>6</sup>Blust (1984c:205) postulates *\*kataman* with some uncertainty, because of Mussau *atamana* beside Emira *atama*. The Molima example does not settle the question, because in that language (which often preserves POC\*-C) *-na* can be a suffix.

were built on the ground, even in coastal areas. Where they are on piles, they are sometimes entered not by a true ladder, but by a slanting log (which may be removed at night). Both these and true ladders may be called by a term which can be reconstructed as *\*tete*, presumably derived from PAN *\*taytay* 'bridge, etc.' The comparative evidence indicates that the basic meaning of the term is 'to walk along a log (or something similar, such as a branch)'. This word, sometimes with a prefix, then became the word for 'bridge' or 'ladder/log for entering a house'. Examples of reflexes are: Nggela *tete* 'to cross a stream on a log or bridge, to descend a ladder, not using hands'; Molima *tete* 'to walk along a ridge or branch of tree', *itete* 'bridge'; Kove *tete* 'ladder, log leading into house', *patete* 'to go up a ladder'; Lakalai *vagege* 'bridge'. I doubt if this set should lead us to attribute ladders to PMN culture, but if we do, I suggest that they were logs which may not even have had notches cut in them (compare Blust's PAN *\*SaReZaSan*).

Although it seems clear that the basic meaning of POC *\*panua* was 'occupied place', and reflexes of the term have a variety of meanings from Fijian 'land' to Molima 'house' to Lakalai 'men', this is the only proto-form that can be cited for PMN 'village', the meaning that it has in languages throughout the region. The situation is a little different as regards the term for 'open space in village'. Grace (1969) gives only *\*malaqe* (cf. PPN *\*malaqe* 'meeting place'), together with *\*mwalala*, later written *\*malala*, with a variety of meanings ranging from 'earth' to 'empty' and including 'cleared ground'. There are a few cases in western Melanesia in which probable reflexes of *\*malaqe* mean 'village': Jabem *malaq*, Wedau *melagai*, with cognates in related languages. Reflexes of *\*malala* seem to be somewhat more likely to mean 'village plaza', as with Lakalai *malala*. It seems likely, however, that *\*malala* basically refers to cleared ground of any sort; cf. Manam *malala* 'market place, assembly place', and perhaps Sengseng *-mla* 'in the open'. Possibly *\*malaqe* and *\*malala* come from a single root, but if they did not, then it is impossible to ascertain which of these led to *mala*, the Nggela word for 'place' and the Tami word for 'village'.

### 2.1.3 VEGETABLE FOODS

Apart from an early suggestion by Groube (1971) concerning the Lapita culture, it has been agreed that the first AN-speaking settlers of Melanesia practised horticulture, cultivating a range of root and tree crops together with other plants. Some of these plants they presumably brought with them from the Indo-Malaysian region, whereas others were discovered within Melanesia – probably already being cultivated by NAN speakers. In the case of sugarcane, however, even if it was first cultivated in New Guinea (Barrau 1958), it must have spread westwards before AN speakers came to Melanesia, because terms cognate with POC *\*topu* are found throughout the AN-speaking region. It may be, however, that the phonemic resemblance between the POC word for sugarcane and one for *Saccharum edule* (*\*tampukal*), a cultigen confined to New Guinea and neighbouring islands, may indicate that the latter term is an adaptation derived from the word for 'sugarcane'. I do not, however, know any languages in which the first vowel in the word for sugarcane is *a* (see Chowning 1963).

Recently French-Wright (1983) has verified or reconstructed a considerable number of POC words which have added to our knowledge not only of the plants eaten but also of techniques of cultivation and preparation. Nevertheless, because his terms have been cited without qualification as "associated with gardening" (Pawley and Green 1984:131), some difficulties should be noted. One is that a number of terms may indeed relate to plants but would apply whether or not they are cultivated. Examples include terms for 'mature' and 'ripe' and 'long pole for harvesting fruits'.

Another is that several of the terms (as is of course the case with *\*matuqa* 'mature') have much wider meanings, and although they may be applied to horticultural and related practices, it is not justifiable to assume that these apparent extensions were part of the original meaning. Examples, cited in the same source, and taken from French-Wright (1983), are *\*pale* designating a garden shed (see above), and *\*pata*, which seems to refer to any kind of platform or shelf, not necessarily one for storing food. Present-day reflexes of these terms do not mean that we can infer the existence of garden huts or food storage on platforms, though it is certainly perfectly likely that both of these existed. We also do not know that yams were planted in mounds or staked everywhere, judging from French-Wright's data, which cite only eastern witnesses of his reconstructions. I do not know, however, the data supporting POC *\*ta(m)puki* 'yam mound' in Pawley and Green (1984).

Of the terms unquestionably associated with horticulture, the best attested is certainly *\*quma* 'garden, to garden'. In addition, despite its occasional use for digging up wild plants, a word for 'digging stick' (dibble) is well attested. French-Wright proposes *\*waso*, which is the shape in Arosi and Tanna, but in view of Lakalai *uaro* and Kove *waro*, I think *\*wanso* is more probable. As regards another term, one I originally proposed (Chowning 1963), I would now amend what has come to be the accepted form, *\*upe*, for a taro top to be planted, by postulating a final consonant to account for Molima *uveya*. A probable form is *\*upeq*.

Within Oceania, the manufacture of sago flour characterises the western part of Melanesia – the New Guinea region, according to Barrau, though he also notes that there were traditions in Vanuatu of it having been made in the past (1958:37-38). In some other areas, such as parts of the Solomons, the pith was fed to pigs, but human beings did not eat the food. What Barrau overlooked, however, was the manufacture of sago in Tikopia and Anuta – according to Yen (1973), one of their Melanesian traits, along with betel-chewing. It seems probable, then, that sago flour was once made in nearby parts of the Solomons.

Even if the palms apparently grow wild in Melanesia and may first have been used for food there (Barrau 1958), the fact that the words for sago palm (PMP *\*Rumbia*, POC *\*rampia*), and sago thatch (PAN *\*qatep*, POC *\*qatop*) are cognate in PAN and POC indicates that the ancestors of POC speakers were already familiar with the plant before they arrived in Melanesia. French-Wright (1983) notes that reflexes of the former term are confined to New Guinea and the Bismarck Archipelago. It is unlikely that they invented the process of manufacturing the flour, since it is made by many NAN speakers. I have only one bit of linguistic evidence for a PMN word related to sago manufacture. This is a cognate set of words for sago beater, the implement used to pulverise the pith, in Kove *walu*, Lakalai *ualu*, and Molima *ewanu*. This suggests a form *\*(e)walu* for western Melanesia, since *\*l* is often *n* in Molima (cf. Molima *nima* 'hand').

The word for 'sago' in Tikopia and Anuta is *ota*. Since the sago palm itself is sometimes called by the word for 'thatch', as in Lakalai *hato* and various examples from the Solomons (Chowning 1963), it is possible that this is a metathesised form of the same word. Another possible origin, however, is PPN *\*gota* 'dregs, rubbish'. The Kove cognate *kota* refers only to the pith left behind after sago starch has been extracted. Since PPN has another word, *\*penu*, for the remains left after grated coconut has been squeezed, it is tempting to wonder whether POC *\*(kq)onta* referred originally to sago pith and was applied to other matter after OC speakers moved beyond the area of sago use, or abandoned its use. Note the Fiji doublet *kora* 'refuse of scraped coconut', *kosa* 'dregs of *yaqona*'. Geraghty (1983) mentions PEO *\*gota*, glossed only as 'coconut grated and wrung', and cites only Fijian *kota*, with that meaning.



Sago is an example of a food plant that may be very important in the economy without necessarily being cultivated. It needs to be noted that a number of the fruit and nut trees listed in Pawley and Green (1984) were rarely or never cultivated in particular Melanesian societies; at most, those laying claim to bush trees might clear an area around the base of the trunk. Examples include *Canarium* almond, *Terminalia*, *Pometia*, and often breadfruit. Furthermore, many other wild plants for which POC names have been reconstructed, were also exploited, for food as well as other purposes. Examples include edible fungi: POC *\*koko* (French-Wright 1983) and *\*taliŋa* (the same word as 'ear', from the shape), and wild mango, *\*wai* (Chowning 1963). Important though gardening undoubtedly was in POC society, wherever local conditions permitted so was the exploitation of bush foods, which added immensely to the variety and nutrition derivable from cultigens.

#### 2.1.4 DOMESTIC ANIMALS

The POC word for 'pig', *\*mpoRok*, is derived from PMP *\*beRek*, as Pawley and Green note, and is attested throughout Melanesia. Compare Lakalai *bolo*, Nggela *mbolo*. Pawley and Green do not mention dogs, which are eaten in many parts of Melanesia as well as being used for hunting. Grace-Lincoln reconstruct POC *\*kaun*, and apparently cognate terms are scattered across Melanesia; see Kove *kaua*, Nggela *kau*. Nevertheless, partly because of the diversity of terms for 'dog' and also because of the lack of archaeological evidence for their antiquity, Hudson (1989) has recently proposed that dogs cannot be attributed to POC, and that the forms apparently reflecting POC *\*kaun* or *\*nggaun* are onomatopoeic. Her evidence looks convincing, but it is necessary to keep in mind the local diversity of terms for starch staples that no one hesitates to attribute to POC (Chowning 1963). I still suspect that the wide distribution in the west of words beginning with *kau*- may reflect a term attributable at least to western Melanesian.

Pawley and Green have been curiously reluctant to reconstruct a POC term for 'fowl/chicken', saying of terms like Bugotu *kokorako*, Motu *kokoroku*, and Tolai *kakaruk* that "apart from irregularities in the sound correspondences, the onomatopoeic nature of these forms reduces confidence in their cognation" (1984:130). Adding Lakalai *kureko* (pl. *kukureko*) to the list in no way alleviates the problem of the vowels, but I think that the correspondence in the consonants goes well beyond anything that can be attributed to onomatopoeia, and that it is necessary to reconstruct forms like *\*kVkvRvkv* for western Melanesia. Cognates extend only into Guadalcanal, however, being replaced further south and west by other terms, such as those derived from Proto Malaitan *\*kua*. Green (1979b:37) notes that chicken (as well as pig) bones occur in Lapita sites, including Watom, off New Britain, but it may nevertheless be worth mentioning, as regards these varying terms, that chickens were not found even throughout New Britain at the time of first European contact. (The Sengseng lacked them, saying that pythons kill any they bring in, and most, though not all, Kove speakers say they are a recent introduction.) None of the three PPN terms for these three creatures have certain cognates in western Melanesia.<sup>7</sup>

#### 2.1.5 FISHING AND HUNTING

My own data add nothing to the relatively rich vocabulary already reconstructed for POC fishing techniques (see Pawley and Green 1984:129; Blust 1984:64). Probably other terms will be

<sup>7</sup>I have recorded elsewhere (Chowning 1987) my objections to the suggestion that reflexes of the PPN word for 'dog' are found in NAN languages of New Guinea.



reconstructed; for example, I should expect one that differentiates a multi-pronged fish spear from the kind used for hunting and warfare (see below). Here I wish to say more about an area that has been comparatively neglected, the hunting and trapping of non-marine animals. The reasons for the neglect probably reflect both the degree of attention received by eastern Oceania, with its scanty land fauna, and the tendency to equate Lapita with POC. Since Lapita sites have usually been littoral, the animal remains found in them have been largely marine. Green (1979b:37) concludes that hunting "was never an important part of the Lapita economy". Even if Lapita can be equated with POC, this conclusion may reflect a bias in the location of sites discovered and excavated to date (see Spriggs 1985). The large number of names of trees and plants, many wild, that can be reconstructed for POC – by no means exhausted by lists in Chowning (1963), French-Wright (1983) and Pawley and Green (1984) – indicates that the people were familiar with a forest environment away from the seashore. They must in any case have exploited it to get material for houses and canoes, and it would be remarkable if they ignored the animal resources. At the very least, they would have had to take measures to control the depredations of feral pigs which were surely present on the larger islands even before AN speakers arrived (Bulmer 1982). (Green assumes that the pig bones found in several Lapita sites were from domestic pigs, but this need not have been true everywhere.)

The most common hunting method recorded in the ethnographic literature required only dogs and spears (and, in grasslands, sometimes fire). If dogs were indeed absent, it is still likely that spears were used for hunting as well as fishing (and warfare). One POC term for 'spear' was originally noted as pan-Melanesian by both Codrington (1885) and Ray (1907). This is reconstructed for POC as *\*io* 'spear', and it appears in that shape in most languages. There are a few exceptions, however; Kove *iro*, Gitua *izo*, Mekeo *iso*, and Wogeo *iwo* all point to the presence of a medial consonant. Unfortunately nothing in Ross's list of reflexes for POC consonants makes it possible to decide what this one should have been. I should add that in Bwaidoga and a few languages of Central Papua the word is *gio* or *Giyo*, but these are languages which sometimes have that initial consonant in words that in POC almost surely began with a vowel. If there was an initial consonant, it was probably *\*q*. For the present, I suggest the reconstruction POC *\*(q)iCo* 'spear'. In addition to this, Blust has reconstructed a PMP term *\*saet* 'spear, to spear' but cites no Melanesian witnesses, nor have I noticed any.

Terms for 'hunting bow' and 'to shoot', dating back at least to PMP, are well attested in Melanesia (see Pawley and Green 1984:131), and Blust has added a term for 'bowstring' in PMP *\*deles*, with reflexes in Sa'a and Arosi (Blust 1980). By contrast, the term for 'bowstring' in Grace-Lincoln, *\*uka*, seems to be confined to the east.

Pigs and other game such as cassowaries were also caught in snares and traps. I have yet to reconstruct a PMN word for 'snare', though I had hoped that the resemblance between Sengseng *sik* and Molima *sikwa* would lead me to other cognates. Blust, however, has reconstructed PAN *\*taqen* 'to set a trap' (1984). In view of Lakalai *taho* 'to set a snare', I feel justified in reconstructing POC *\*taqon*.

Sticky sap was often used as birdlime, but while terms in some languages (e.g. Lakalai *bulubulu*) may derive from PAN *\*bulit* / *\*puli*(B) 'stick, glue, paste', the evidence does not justify ascribing birdlime as a primary meaning of the term.

Pawley and Green list POC *\*su(n)ja*, from PMP *\*suja* 'sharpened stake set in ground to stop or wound animals or enemies'. I assume that Molima *suna*, 'sharpened stakes set where a pig jumps (as over a garden stile)' reflects this. On the other hand, Kove *sokasoka*, for the same apparatus, derives

from POC *\*soka* 'stab', which in Grace-Lincoln has reflexes only in Fiji, Rotuma, and Polynesia. This Kove form suggests that the term can be attributed to PMN.

I hope that in time it may be possible to reconstruct other terms relevant to this topic, such as words for 'pig net' and 'bird net'. (In addition, I would expect to add some words for weapons used primarily if not exclusively in warfare, such as clubs.)<sup>8</sup> At this point I argue only that the statement in Pawley and Green (1984:128-130) that "POC speakers had an economy based jointly on gardening and fishing" probably seriously underestimates the degree to which they exploited the resources of the bush, both fauna and flora.

### 2.1.6 FOOD PREPARATION

A few terms already reconstructed for POC, although they often have wider meanings, are so often associated with aspects of food preparation throughout Melanesia that it is probably safe to attribute those meanings to them. These include *\*kosok* 'to husk a coconut, pointed stick used for doing this'; *\*kweli* 'scrape a coconut' / *\*kori* 'scrape, shave, peel'; *\*supi* 'to peel, as taro'; *\*kasi* 'kind of bivalve; to scrape (coconut, char from roasted food) with the edge of a shell'; and perhaps *\*poRo* 'to squeeze, wring out juice, as coconut into food'. Coconut oil itself is often called by a term like *mona*, especially in western Melanesia, by contrast with the east, where the term used in Polynesia, *lolo*, is more common (but see Nggela *mona*, cited in Blust 1984c:205). The former term, from POC *\*moñak* 'fat, grease', (including that of animal origin), often has a secondary meaning 'sweet, pleasant-tasting', as with Manam *monamona*. A more specific meaning is also found in a few western languages: pudding of cut-up starchy vegetables boiled in coconut cream. Examples are Kove, Molima, and Kilivila *mona*, and probably Motu *mone* 'cakes of sago, taro, etc. boiled in leaves'. I had originally postulated this gloss on the basis of the shared forms in Kove and Molima, and just learned that Blust had done the same though more tentatively, because of Mussau *mona* 'pounded taro with coconut cream' (Blust 1984c:193, 205). He cites a Sa'a term that would justify attributing this meaning to POC. Unrelated words in many other languages are glossed 'pudding'; this particular term seems to have a limited distribution.

Linguistic, archaeological and ethnographic evidence agree in justifying the reconstruction of at least five terms, two nouns and three verbs, related to cooking methods. These distinguish between boiling in a clay pot, steaming between heated stones, and roasting directly on the fire. As regards the first, the distribution of potsherds in sites, some of which predate Lapita, has led to general agreement that the early speakers of OC used pottery, even though many of their descendants did not. Where it did exist, however, a word derivative of *\*kudon* is so common that the term must be ascribed to POC. It is usually a noun, so that the identical verb, meaning 'to boil in a pot' (e.g. Kove *ulo*) is presumably a secondary formation. In addition, Milke (1965) reconstructed a verb for 'boil', *\*nansu*, for PNGA alone. I criticised part of the cognate set involved (Chowning 1973), but Blust (1980:115) assigned *\*nansu* to PAN and cited cognates of Motu *nadu* in other Melanesian languages. Teop of Bougainville has *nahunahu*. Nevertheless, reflexes of this form are so rare that I expect that eventually it will be possible to reconstruct another PMN/POC word, possibly simply meaning 'to cook'.

<sup>8</sup>Although shaped sling stones have been identified in Lapita sites (Green 1979b:39), the POC word *\*maka*, which has 'sling' among its glosses, seems to mean that only in Polynesia.

The term 'earth oven' is a misnomer in many parts of Melanesia, including Lakalai, Kove and Sengseng, where the whole process is completely above ground. It seems better to accept the gloss 'stone oven', even though the evidence at Lapita sites indicates at least some excavation. The method of cooking is essentially the same in all cases, except that extra water is not necessarily added to foods that are juicy in themselves or well wrapped in leaves. The POC term was given as *\*qumu* in Grace (1969), but Blust amended it to *\*qumun*, noting that the final consonant in Molima *'umula* is anomalous. Sometimes, as in Lakalai *humu*, the same word is used for the oven and for cooking in it, but there is abundant evidence for reconstructing POC *\*taqo* for the process. This term was originally reconstructed for PEO, and is so listed in Grace (1969). I have not been able to locate it in Grace-Lincoln. Nevertheless, it is present in the west with the same meaning, as in Kove *tatao*.

Probably the most common form of cooking in Melanesia involves simply laying the food on an open fire, or hot embers, whether wrapped in bark or leaves or unwrapped. Throughout Melanesia and Polynesia, the same term, from POC *\*tunu*, was used for this process and also for 'set alight, burn'. (Unrelated terms in some languages, such as *gabu* in Central Papua and Milne Bay, have the same range of meaning.) Grace-Lincoln give as the only western cognate Motu *tunu*, 'bake pottery', but Tolai *tun* is one of many western examples that make it possible to say that this was a pan-Melanesian term for roasting, in addition to POC/PMN *\*maRi*.

Other terms for 'cooking' have also been proposed, such as Milke's *\*kasaj*, and there exist a number of terms (glossed 'mix', 'stir', 'knead' etc.) which probably refer to processes of food preparation. I have not to date been able to confirm a wide Melanesian distribution of any of these, or of a word for smoking food, reconstructed by Blust as *\*tapa*. Blust's data (pers. comm.) connect Roviana *tava* with numerous non-OC forms, and therefore do indeed justify assigning the term to POC. Milke's few examples from which he reconstructed *\*sagan* do look cognate, but the meaning of the term is uncertain. Nevertheless, I suspect that in time various other terms related to food preparation will be found to be distributed throughout Melanesia.

The one process that is distinctively eastern is that of preserving vegetable food by fermentation. Green attributes this process to the Lapita stage, but if it was present so early, it is curious that no trace of the process remains in the west (see Barrau 1958; Yen 1973; Pollock 1984 on the distribution of the processes). In the east, two different terms have been used to describe the process and/or the product. One derives from POC *\*ma-asin/\*maqasin* and another from POC *\*mada*. Both of these were originally attributed to PEO. The former term generally, and surely originally, meant 'salty', but derived meanings ranging from 'stinging' (like salt water in a wound) as in Lakalai, to 'tasty' and 'sweet' as in Sengseng and Cristobal-Malaita languages (Blust 1984b). The other, POC *\*mada(d,R)*, is glossed in Grace-Lincoln 'soft, ripe, fermented', but I suspect that the basic meaning was 'over-ripe', as in Tolai *madar*. It is only in Melanesia that the term for fermented food seems to derive from *\*ma-asin* (see terms in Pollock 1984). In the Solomons, a process of fermenting taro by adding coconut cream mixed with salt water seems to be a separate development, and the product sometimes called by the same term that is used for 'coconut cream' (eastern *lolo*). Despite the archaeological pits, we lack linguistic evidence for attributing fermentation to any part of western Melanesia – nor, as noted, is it recorded in the ethnographies.

The one implement apart from clay pots that can reliably be attributed to POC is tongs for handling cooking food. Although other terms exist, the evidence for pan-Melanesian distribution most strongly supports *\*kapit* 'tongs', (so defined in Grace 1969).

## 2.1.7 CONTAINERS

The words for 'water bottle' and 'lime container' tend either to be the words for 'water' and 'lime' or a general name for the material – coconut shell, bamboo, gourd. I have not been able to reconstruct a PMN term for these two types of containers. The best data exist, interestingly, for two artefacts that are not found everywhere in Melanesia, clay pots and wooden bowls, and for a third one that differs in its material and shape from one society to the next. I have nothing to add to the data already presented to justify POC *\*kudon* 'clay pot'. Wooden bowls vary in shape from the deep elaborately carved Tami (or 'Siassi') bowls so widely traded in north-east New Guinea and across the Vitiaz Straits, to what sound like shallower 'dishes' in the Solomons, but cognate terms, derivable from POC *\*tampida(ŋ)*, are found from the north coast of New Guinea (Ross 1977) to Vanuatu (Merlav *taber*, cited in Pawley 1972:114). The term *\*taga*, found throughout Melanesia, seems to refer to either a bag or a small basket but typically designates the kind used for carrying personal possessions such as betel-chewing equipment. I have not, however, been able to find clearcut evidence that words for larger baskets, cognate with PPN *\*kato* and *\*kete*, were used in the west.

I am also uncertain about the evidence for another term, *\*ipu* 'container for liquid'. The only non-Polynesian form cited in Grace-Lincoln is Ulawa *ipu* 'hollow in tree holding water', and the only possible western cognate that I can cite is Lakalai *piu* 'drinking coconut'. Several other words in Lakalai that reflect POC *\*iCu* show this kind of metathesis; cf. *liu* 'drink', *kiu* 'tail'. Nevertheless, since neither the Lakalai nor the Ulawa terms designate what the Polynesian ones do, I am reluctant to suggest that this was actually a POC term for a kind of container for liquid.

A word related to containers that has not, so far as I know, been reconstructed is, for the western region *\*quntu* 'to carry on the head'. In Lakalai *hugu* and Kove *uru* refer specifically to the way that women carry burdens, including baskets and, in Kove, wooden bowls (*tavila*), though the Lakalai also use the term for a man's wearing of a mask to cover the head. The Motu word, obviously cognate, is *udu* 'to carry a child astraddle the neck'. I have not found cognates in the east. By contrast, Sengseng *sun* 'to carry on the head', may derive from PAN *\*suqun*.

## 2.1.8 MATS, CORDAGE AND NETS

In most precontact Melanesian societies, pandanus mats were formed of strips sewn together rather than plaited. Plaited mats, traditional in Polynesia, were then introduced into many parts of Melanesia by Polynesian missionaries. With one end sewn together, mats often doubled as raincoats, and sails were sometimes made in the same way, though in other areas sails were plaited (see Haddon and Hornell 1975). Although often the word for 'pandanus' was also used for all of these artefacts, the evidence for one POC word is well known. Its pan-Melanesian distribution was pointed out by both Codrington and Ray, and it has also been attributed to PPN and PPH. I emend Grace's POC *\*qempa* to take account of Molima 'ebana'. The PPH form is *\*hempas*, which raises questions about the final consonant; Molima /n/ reflects *\*n*, *\*l* and *\*ŋ*. I propose only *\*qempa(C)*.

In 1963, I reconstructed a 'Proto Melanesian' term for pandanus in addition to *\*pandan*. This was *\*moi*, and my evidence then included 'Admiralties' *muoi*,<sup>9</sup> Lakalai *moe*, and Wedau *moi*. To these I can now add Kove *moe* (contrasting with *moi* 'taro'), *omoi* in several West New Britain languages (Thurston 1987:130), and Kilivila *moi*. In most of these languages the word also designates sleeping mats. Apart from uncertainty about the final vowel, we also need to note Bola *moke* 'pandanus, rain

<sup>9</sup>Blust (pers. comm.) tells me that this term cannot be cognate with the others.

cape' and Mandok *moki* 'mat'. They point to a proto-form *\*mok(e,i)* 'pandanus, sleeping mat'. There is no reason to assume a connection with POC *\*mose* 'sleep'; *\*s* would not become /k/ or /θ/ in these languages. Apart from the dubious cognacy of Gedaged *moi* 'dull, inactive', reflexes of *\*mose* seem to be confined to eastern Oceania. By contrast, the words for 'pandanus, sleeping mat of pandanus' that originally led me to reconstruct *\*moi* are only attested in the west, with one probable exception: Kwaio *mode* 'pandanus, mat or umbrella made of pandanus'. These are also a rare example of terms that connect the Admiralties with other parts of western Melanesia. The medial consonant is a problem, but it seems likely that we should ascribe *\*mok(e,i)* to PMN.<sup>10</sup> Another very widespread term in eastern Oceania derives from *\*pola*, reconstructed in Grace-Lincoln for POC and defined 'coconut thatch'. It is clear from the meanings of the cognate forms outside Polynesia, however, that it designates coconut leaves woven together for any purpose, including mats and bags. See for example, Efate *bora* 'coconut leaf, or basket made of it, or plaited for thatching houses' (Macdonald 1907). Again, this seems to be an eastern term, though another and possibly related term *\*pola* 'spread out, lay down, as a mat' has possible cognates in western Melanesia (see Lakalai *bolabola* 'flat'; Kove *pola* 'open out'). But since 'mat' does not seem to be the primary meaning of the noun, I suspect that the verb is unrelated.

The sewing together of leaves, and similar acts, are often expressed in Melanesia by reflexes of POC *\*saqit*. The objects sewn are most often pandanus mats, as in Lakalai (where the word *sahi* is also used for 'sewing' thatch), but the Sengseng, who lack mats, use the reflex *sihit* both for the sewing of areca palm spathe pouches for lime powder, and for the barricading of doors by interlacing vines across them. There is also abundant evidence for the reconstruction of POC *\*nsaRum* 'needle' (PAN *\*zaRum*). (This was *\*nsaRu* in earlier POC reconstructions, but the final consonant is necessary to explain Kove *salumu* and Molima *saima*.) In some Melanesian languages, words for 'sew' have other origins, such as POC *\*tuRia*, but it seems clear that, as Grace-Lincoln indicate, the basic meaning of this is 'to string on a line'. I am sceptical about some of the other POC terms that have been reconstructed for 'sew', such as *\*susud*, reflected in Rotuman and Sa'a *susu* and Mota *susur*. This seems suspiciously close to POC *\*sudu* 'enter, insert'. It may be that in PMN, there was only one term for 'sew', though others developed out of related terms in various parts of the region.

Thread was sometimes only unmodified plant fibre, as from the inner bark of hibiscus, or thin vine. Thicker cord was, however, often composed of several fibres rolled together. Grace-Lincoln contains three reconstructions which resemble each other in form and meaning: *\*pidi* 'twist, bind around, plait, braid'; *pili* 'plait, wind around'; and *\*widi* 'turn, revolve, twist around, twist, bore'. The two PPN forms *\*firi* 'plait, braid' and *\*wili* 'twist, bore' suggest that indeed two proto-forms were involved, but if so, the reflexes in some languages seem to be identical (cf. Maori *whiri* 'twist, plait'). It is tempting to suggest that a term which had been used for the manufacture of cordage by twisting (from a word that originally meant 'to twist, to wind (or bind) around') was then extended to include a different method of manufacture by braiding.

Through Melanesia the word for 'vine' and rope' are the same, derived from POC *\*waRos*. Presumably this reflects the fact that often strong vines (or rattan, POC *\*quwe*) were used for binding without being modified. The evidence is also good for a term *\*tali* 'cord, line', typically used for

<sup>10</sup>Tryon and Hackman do not derive Kwaio *d* from any POC consonant. It may be that the medial consonant is the result of word tabu, which frequently produces phonemic change in Kwaio words (Keesing 1975). The dictionary shows examples of the opposite shift, as in the doublets *oka/oda* 'eat raw food', when the historical evidence shows that the latter is the expected reflex.



something smaller than a rope. In some languages, reflexes of it mean 'sew', but this is probably a secondary meaning (as Grace-Lincoln seem to agree).

Another term that looks like a good candidate for PMN is *\*loi*. In both Kove and Shortlands the identical form is the term for manufacturing thread by rolling the fibres on the thigh. In Lakalai, where that process is described by a verb derived from the word for 'rope' (Lakalai *ualo*), *loi* means 'to rub between the hands'. But outside western Melanesia, see also Kwaio *loi(a)* – 'weave a net, knit'; Efate *lolo* 'thread', *loloa-si* 'manufacture thread' (apparently derived from *loa* 'rub', however); and outside Melanesia, Gilbertese *ro* 'cord'; Tuvalu *lö* 'binding, sennit'; Maori *roi/roki* 'secured, tied up, knot, bind'. Within western Melanesia there are also Tolai *loe* 'to coil', *loloi* 'roll or coil' (both terms referring to how strips of shell money are made into coils), Manam *lolori*, (pl. *lolo*) 'to bind or knot grass skirts'. If none of these represents a shortening of *\*waRos*, there may still be reflexes of at least two proto-forms here, one meaning 'tie up' and the other 'roll (fibre?) under the palm of the hand'. For the present, I shall suggest that *\*loi* 'roll under the palm, make cord' is justified at least for western Melanesia.

I have a small amount of evidence to suggest the reconstruction of another term used for the process of braiding larger ropes, such as those used on canoes. This is *\*mwali*, reflected by Lakalai *mali*, Molima *mwali*, and also, I assume, by Manam *moli* 'plait'. I have not, however, found reflexes in other languages. Because some writers use 'plait' for 'braid' (as cordage) and others for the process of mat-making (otherwise 'weave'), it is often difficult to know which process is being described. Nevertheless, these three examples should suffice to establish this term for PWMN. Various other terms have also been proposed, including *\*(q)a(ñ)am* 'loosely braided' (Blust 1978:5).

Although POC words for 'fishnet' have been reconstructed by others (see references in section 2.1.5), most terms relate to the finished products rather than to techniques of manufacture. POC *\*sika* was reconstructed by Milke (1961) for 'netting needle', reflected in PPN and Fijian (*sika ni lawa*), and also in Gedaged *siwali*. Grace-Lincoln follow Milke in including as cognate Sa'a *sike* 'thorn', but a thorn is not much like the shuttle that has been glossed as needle. Another probable cognate is Molima *sikwana/si'wana*, which is also the verb for manufacturing nets. The Gedaged and Molima terms suggest a final consonant for the proto-form. In addition there exist two other terms which are certainly cognate with each other, Dobu *siyona* and Kove *sione*. In Kove, and in some other Milne Bay languages, though not usually in Dobu, labialised stops preceding a shift the *a* to *o*: compare Lakalai *pati*; Manam *pwati*; Kove *popoti* 'float'. If all of these terms for 'net gauge' are cognate, they suggest a proto-form like *\*sikwan*, though it would be helpful to have other cases in which *\*k* disappeared after affecting the following vowel. See also Motu *diva*; according to the reflexes in Ross (1988), this should derive from POC *\*siwa*.

In Kove and Molima there exist cognate terms for 'net gauge': Kove *mata*; Molima *matana*. Since *mata* is the word for 'mesh' in many OC languages (see discussion of 'doorway' above), I assume that the name of the tool is a secondary development, and perhaps independent in these two languages. I have not yet found any other examples. Another term for net gauge, *\*qapa*, has no recorded reflex in the west.

### 2.1.9 WATERCRAFT

Pawley and Green list a considerable number of POC words associated with 'sailing', but do not include all of those to be found in Grace (1969) and Grace-Lincoln. My own data on this topic are not very full, so that all I can do here is slightly amend either the shapes or the glosses of some proto-



forms reconstructed by others. Blust defines PMP *\*dakit* as 'join along the length, raft'. For Oceania he mentions only Motu *rai-a* 'prepare a raft for the sea' (Blust 1986). Lakalai *lage*, Kove *laGe* 'raft' presumably derive from this proto-form but leave the precise shape of the POC form uncertain.

Grace-Lincoln only gloss *\*quliŋ* 'to steer', but in a variety of languages the cognate terms also mean 'steering paddle' or 'rudder'. This is also the case with Molima *kuliga*, and I suggest that this meaning should be one of the primary ones. Indeed, Pawley says as much (1981:287). Similarly, my own data indicate that in a number of languages, including several in Polynesia, reflexes of POC *\*tokon* 'pole, stick' not only designate walking sticks but also punting poles, and so may justify our attributing these implements to POC. (Kove, however, contains a doublet: *atoko* 'walking stick', *toto* 'to punt a canoe'.)

#### 2.1.10 MUSICAL INSTRUMENTS AND GAMES

The one implement, only marginally musical, that can certainly be ascribed to POC (and PAN) is the conch-shell (triton shell) trumpet. Cognate terms derivable from POC *\*tapuRi* are found throughout Oceania, and although Grace-Lincoln do not include 'trumpet' in the gloss, Wurm and Wilson (1975) do so with respect to the PAN form. The use is so frequently mentioned that I consider it safe to associate it with these shells (which are not always the same kind). My own data did not permit me to reconstruct any other POC terms with certainty, but Blust (pers. comm.) tells me that the Fijian term for 'nose-flute' has cognates in several Indonesian languages, so that "POC speakers must have had the nose-flute".

Other more purely musical wind instruments are also very widely distributed, particularly flutes and Pan's pipes. I have not, however, been able to reconstruct terms apart from those which reflect words for 'bamboo', such as POC *\*kaundu* (Blust), and those which reflect one or more POC words beginning with *\*pu-*, which mean 'to blow'. Grace-Lincoln contains three such terms: *\*pu* 'explosive noise, conch trumpet'; *\*puput* 'to blow' (the same in PAN); and *\*pus-i* 'puff, blow', also from PAN. I suspect that other terms should be reconstructed, but the 'blow' sense of the first syllable is obvious. An example of the problems of disentangling related meanings can be seen in Lakalai: *pupu* 'to blow up' (as a pig bladder, for a balloon, or blowfish, both called by terms derived from this); *pusu* 'to blow, as conch shell'; *e pususu* 'small wild bamboo and articles made of it, including Pan's pipes and flute', 'flutemouth' (fish); *puratete* 'of a masker, to produce a particular sound by blowing through tense lips'; *pusese* 'shrill sound produced by masker through pursed lips'; *pututu* 'to blow pipes or whistle', *e pututu* 'whistle made of split bamboo'.

In western Melanesia, two types of percussion instruments are widely distributed, among NAN speakers as well as speakers of OC languages. One is the hourglass drum with a skin head. To date I have not been able to reconstruct a name for it that extends throughout the area. Along the entire north coast of New Britain we find cognates of Tolai *kudu*, the origin of the Tok Pisin name; examples include Lakalai *kude*, Kove and Kilenge *kure* (see below). Mota *kore*, though defined as 'slit gong', is the only possible cognate I have found to the east or south of this region. On the north coast of New Guinea and in some parts of West New Britain, the name is suspiciously close to the word for 'monitor lizard', whose skin is often used as a drumhead (see Thurston 1987:127), though in Sengseng they differ minimally: *pahiyo* 'drum', *apahiya* 'lizard'. Terms in Milne Bay and Central Papua differ from both of these, and from each other. I do not have the data to indicate whether any borrowing from NAN languages may be involved.

Niles (1983) has made a study of the distribution of slit gongs and terms for them, and points out that in New Guinea slit gongs are limited to the central north coast and regions inland from it, including the whole of the main stream of the Sepik. They do not occur in the region formerly called Papua, nor on the western end of the island, but they are found throughout the Bismarck Archipelago, in the Solomons, and in Vanuatu. The names for them in PNG show many cognate forms, though not quite so many as Niles suggests. For example, Kilenge *na kure* is surely cognate with terms for hourglass drum found elsewhere; the Kove often use that term for both instruments. Again the most common terms are cognate with Tolai *garamut*, the source of the Tok Pisin term. Terms like this, including the final consonant, are found throughout New Ireland, up into Mussau-Emira (*galamutu*), and across into Halia of northern Bougainville (*garamuts*). In West New Britain, the final consonant disappears, so that the term is *galamo* in Lakalai and its closest relatives, and then further west the first vowel changes, producing Kove *Gilamo* and related terms in New Guinea (Kairiru *giram*, Manam *giramo*). Terms in the Admiralties, such as Bipi *drami*, could be cognate only if they had changed considerably from a proto-form which, for western Melanesia, must have been something like *\*ŋk(a,i)ram(o,u)t*. (Blust, however, postulates POC *\*lali* 'slit gong' based on cognates in the Admiralties, Fiji and Polynesia.) Other terms recorded farther south, in the Solomons and Vanuatu, are very different.

Niles notes that some have suggested derivation from PAN *\*giriŋ* 'bell, gong' (which has also been given as the source for POC *\*kidi* 'ring a bell, beat a drum'). If so, there have been innovations in form and, I assume, meaning; though slit gongs do exist in Indonesia, I would expect reflexes of *giriŋ* to refer to metal gongs of a very different shape. Because of the odd distribution of the instruments, Niles has suggested that they are correlated in western Melanesia with a pre-Lapita settlement of Oceanic speakers (1983:98).

I do not know of any reconstructed term for POC that means 'to sing'.<sup>11</sup> Terms differ widely from one area to another; the few observable cognates are very limited geographically, even when they cut across boundaries postulated for major subgroups (as Lakalai *bau*, Kove *vou*, Gitua *bwau*).

The term for the famous Polynesian dart game, and for the dart, *\*tika* in PPN, has, according to Grace-Lincoln, cognates only in Fiji, Rotuman and Nggela. I have been able to find no others, and suggest that this is another term that should be ascribed to a lower level than POC. The only other game for which I have comparative data is string figures. These are made throughout Oceania, and the relation between Kove *wawaiŋa* and EP *\*fa(q)i* looks suggestive. The difficulty is that the words are very close to those for 'weave' (e.g. Kove *wai*, *wawai* 'to make a basket, plait leaves for a door') and so may represent independent invention.

## 2.2 RELIGION

Several terms can undoubtedly be attributed to PMN. One, which derives from PAN, is *\*qanitu* 'ghost (of the dead), evil spirit'. Melanesian examples include Kove *anitu*, Lakalai (*la*) *hitu*. (I assume that in Lakalai metathesis was followed by assimilation of the first syllable to the article *la*, from POC *\*na*.) In the Shortlands, the term is *nitu*. It seems probable that the name of a major god or gods along the Rai Coast of New Guinea and also among the Mengen of East New Britain, Anutu or Anut, is a variant of this. Blust, who overlooked my mention of the Kove form (Chowning 1973:198), mentions Wuvulu *aniqu* 'ghost' and Wayan *anitu*, the latter supporting his conclusion,

<sup>11</sup>Blust tells me that he has reconstructed such a term, of which Molima *wali* is a reflex.

with which of course I agree, that "the sporadic loss of the nasal in PPN *\*aitu* 'ghost, spirit' postdated the breakup of Proto-Oceanic" (Blust 1978:10). Pawley (1985) also discusses a POC form *\*taumate*, literally 'dead person'. Although terms derived from this may mean 'ghost', they often just mean 'corpse'. I am dubious about whether this term originally pertained to the religious realm.

Another term is *\*qanunu* 'soul, shadow, reflection'. Nowadays reflexes are usually the word for 'photograph'. A reduced form of this, *\*nunu*, had previously been reconstructed (Grace-Lincoln), reflecting the form found from the Shortlands south (compare loss of the first syllable in Shortlands *nitu* with *nunu*); it is reconstructed for Proto Malaita by Levy and Smith (1969). The short form is also recorded for Ulau-Suain, on the New Guinea coast. But the forms that require an initial syllable include Lakalai *halulu*, Molima *'anunu*, Bwaidoga and Kove *anunu*, and Manam *anunuka* 'shadow'. This last may, of course, also point to the presence of a final consonant, but since it is not reflected in either Kove or Molima, both of which often preserve POC final consonants, I have not postulated it. Blust (1978:6) reconstructed *\*(q∅)anunu*. I am not sure whether reflexes of the term are found south of the Solomons, but a few terms from Vanuatu languages look plausible. These include Efate *anu-na* (and Macdonald also cites 'Malekula' *nunu*) and Kwamera *nanumu*- 'spirit, ghost, shadow, image, reflection'. In Kove and Sengseng, reflexes of *\*qanunu* are also the term for 'dream', used as in Kove *anunu-Gu la* 'my *anunu* goes = I dream', but since there is good evidence for another POC word for 'dream', *\*nipi*, this may just be a localised idiom. Like Grace-Lincoln, I am inclined to connect *\*qanunu*, at least as regards the meaning 'shadow', with PAN *\*ali(n,n)u* 'shadow'.

Blust has also postulated PMP *\*qatuan* 'deity', connecting WMP terms for 'lord', such as Malay *tuhan*, with Polynesian terms which derive from PPN *\*qatua* 'deity'. The link between these differing forms is Emira *otuana* 'spirits' (Blust 1984:41). (In Mussau, however, the same word means 'snake'; Blust 1984c.) I am not very happy about the postulated connection or the gloss. It seems surprising that no other Melanesian cognates have been identified (but see below). If they are located, I should not expect the term to mean 'deity'; one of the differences between Melanesian and Polynesian religions is the scarcity, and often the complete absence, in Melanesian systems, of any beings powerful enough to be given that label (see Chowning 1986). If the Emira term is indeed a cognate, I would expect the POC term originally to have designated spirits that were not of human origin, with 'deity' a special development of PPN. It might however be worth taking a look at a term *\*qadua* 'soul' postulated by Capell (1943) and included in Grace-Lincoln. All of the reflexes come from the Milne Bay region (e.g. Molima *yalyaluwa*), and so do not justify reconstruction for POC. Furthermore, tempting though it might be to suggest a connection between *\*qadua* and Emira *otuana*, the likelihood of cognacy is small, judging from the reflexes that Ross gives for POC consonants in Emira.

A completely uncontroversial POC term is *\*tampu*. Reflexes of it are found throughout Melanesia, as well as in Polynesia. My data suggest that the greatest complexities of meaning are found in the eastern region (see Keesing 1982), but this impression may only reflect the rather limited range of meaning that reflexes have in languages I know. For example, in Molima *tabu-* (with a suffixed possessive) refers only to food tabus, and in Kove, also with a suffixed possessive, *tavu-* refers to tabus on some foods, on saying the name of certain affines, and on other types of behaviour. (In both cases, the tabus relate to kin ties.) Nevertheless the comparative data seem to justify the various glosses given in Grace-Lincoln: 'sacred, forbidden, ritual restriction protected by supernatural sanction'.

The situation is somewhat different as regards *mana* terms. Blust has corrected the *\*manan* form in Grace (1969) to *\*mana(q∅)*, though without altering the earlier definition 'power, wind'. I am

dubious about whether a single form is represented here. Only in Milne Bay are there words like *mana* meaning 'wind', and it was Capell who argued that they are related to words meaning 'power' elsewhere. Furthermore, with a single exception not known to Capell, *mana* forms in Milne Bay do not refer directly to power. Because Blust (1978:11) cites it in support of his alteration of the final consonant postulated by Grace following Capell, it should be pointed out that Dobu *bomana* is properly *boma-na*, with the final syllable representing a suffixed third person singular possessive, a fact known to Capell as well as to Fortune (1932:233), who defines the term *boma-gu* as 'my sacred prohibition' as opposed to Capell's 'supernatural power'. Both of these men did relate the term to *mana* as it is used further east, and assumed that the Dobuans had mis-analysed the term, but I do not think we are justified in accepting this assumption. In fact, in an intensive search of the available material, Keesing found only one possible example of *mana* with the meaning in the west, ironically in a Milne Bay language, that of Tubetube (Keesing 1984). Here, however, the shape is anomalous, being written *naManaMa*, and the meaning contains no reference to the supernatural (Keesing 1984:147). Keesing argues throughout his paper that POC *\*mana* did not necessarily refer to the supernatural. But apart from this Tubetube case, reflexes of *\*mana*, often with religious connotations, are very much confined to eastern Melanesia, together with Polynesia. The question of the association of leadership with *mana* is discussed under 2.3.

## 2.3 SOCIAL ORGANISATION

The only two aspects that have been dealt with in any detail are leadership and kinship. To these I shall add, very briefly, only one other: gender relations.

### 2.3.1 LEADERSHIP

A general discussion of the topic of linguistic evidence for POC leadership patterns began only recently, with a paper by Pawley (1982). In it he challenged the common assumption, made famous in a paper by Sahlins (1963), that Melanesia, characterised by achieved leadership, the so-called Big Man system, contrasted with Polynesia, with an institution of chieftainship as a sacred, powerful, ascribed office. It had been pointed out that hereditary leadership was much more common in Melanesia than was commonly assumed (Chowning 1977; Douglas 1979), but a term for 'chief' had been reconstructed only for PPN (*\*?ariki*) (Biggs, Walsh and Waqa 1970). Beginning with an apparent cognate of PPN *\*?ariki* in Arosi which designated the chief's firstborn son and heir, Pawley went on to reconstruct two POC proper nouns, *\*qa-lapa(s)* 'Great One' (originally designating the chief) and *\*qa-diki* 'Little One', designating his successor (*\*qa* marks a personal name or title). *\*qa-diki* then became the Polynesian term for 'chief', *\*qa-lapa(s)* being retained only in various parts of the Solomon Islands, extending as far north as Mono-Alu.

In 1986 Lichtenberk challenged this scenario on several grounds. First, he changed the reconstruction of *\*qa-lapa(s)* to *\*ta-la(m)pat*, but still ended up with a term meaning 'Great One'. He pointed out, however, that such a meaning would not distinguish a chief from a Big Man; reconstructing it gave no decisive evidence for the existence of chieftainship. (The Big Man system is so-called because the titles so often have that literal meaning – see Chowning 1979.) He also suggests that *\*qa-diki* should be *\*qa adiki*, and that the original meaning was 'eldest child', pointing to the importance of primogeniture in POC society. There are, however, some difficulties with this reconstruction. The reflexes of *\*qa adiki* that he cites nowhere mean 'oldest child' (though see below). In Malaitan languages they mean 'unmarried girl, daughter', which Lichtenberk thinks is



(along with the Arosi reflex) likely to be derived from 'oldest child'. I do not find this derivation persuasive. Furthermore, Lichtenberk overlooks cases cited by Codrington (1885:64) in which there are apparent cognates in various places outside the south-east Solomons with the general meaning 'child'. For example, Codrington gives *karik* for Efate; Macdonald has variants *kariki*, *karikiki*. While I could not confirm from Lanyon-Orgill Codrington's suggestion of a cognate form in Tolai ('New Britain'), I was struck by the resemblance to Kove *Galiki*, which designates both a girl in seclusion after her ears have been slit (the house in which she is kept being called *luma Galiki*), and an unmarried heroine in stories. The same term appears in Bola stories and is translated 'Mädchen' (Kroll 1938), but I suspect that this may be a borrowing from Kove. The comparative evidence seems to suggest that the initial consonant was part of the root, and that perhaps the word originally meant 'child (not 'offspring') especially girl'. This seems especially likely if the word indeed incorporates the POC root for 'small', since there is no reason to stress the smallness of a firstborn; cf. Lakalai *guliliki*, Bola *guriki* 'child' (and perhaps Mussau *aliki* 'child' – Blust 1984c). The suggestion that girls in particular were designated would, however, be contradicted by Pawley's reconstruction of a POC term *\*ŋkeni* with this meaning, unless this was a term for 'unmarried woman' (Pawley 1985:101).

There is, however, one piece of evidence to support Lichtenberk's hypothesis. Scaletta (1987), writing of the Bariai (her Kabana), whose language is very close to Kove, says that originally *galiki* designated firstborn children of both sexes, whereas now it is only a personal name for a firstborn daughter. (In Kove, a group of sisters in a story can all be referred to as *Galiki* (pl. *GaliGaliki*) so there is no trace of such a meaning there.) I remain reluctant to assume that the proto-form necessarily designated firstborns. (I also dispute Lichtenberk's assumption that the word was a kinship term, at least outside Polynesia (Pawley 1981:284). In the languages I know it is a title which does not take the suffixed possessive of kinship terms.) I do not dispute that firstborns are often singled out, though not only by AN speakers; compare the Siuai (Oliver 1955). In some languages the term for firstborn is the same as or derived from the word for 'first, precede' which is usually reconstructed as *\*muqa*;<sup>12</sup> cf. Kove *muGaiai* (*muGa* 'precede') and Ambrym *mo*, with both meanings (Paton 1973). Whether or not he has succeeded in reconstructing a word for firstborn, however, I fully agree with Lichtenberk (1986:353) that: "At present there is no convincing linguistic evidence to determine whether leadership in Proto-Oceanic society was of the chiefly or the big-man type".

Nevertheless, in recent years the assumption that POC society had hereditary leadership associated with *mana* has become widely accepted, as by Keesing (1984) and Kirch (1984). Both of these cite Pawley in support. Keesing (1984:15) wrote as follows:

those entitled to lead presumably had to demonstrate, by success in war, skill in leadership and resource management, and proper conduct, that they had the support of the supernatural invisible beings, gods and ancestors on whom life depended. Such success was continuous visible evidence that the leader himself was *mana* or 'had *mana*'.

There are, however, several difficulties about Keesing's interpretation. First, it is at odds with the assumption (made by Kirch as well as by Pawley) that *mana* was tied to seniority of descent so that normally the senior man in the senior line was superior in this respect; he did not have to demonstrate it to achieve office. Second, the assumptions made about the nature of POC religion are not justified

<sup>12</sup>Considerable evidence (e.g. Lakalai *muga*, Kove *muGa*) indicates that at least in the west, a form should be reconstructed like *\*mugka* as an alternative to *\*muqa*. Lincoln (n.d.) discusses the problem as regards Gedaged.

by the cross-cultural evidence. Even where hereditary leadership was found in western Melanesia, it often had nothing to do with gods or ancestors; to the extent that religion was involved, what mattered was control of powerful magic, including sorcery. This was the case in, for example, the Trobriands, Mekeo and Kove. Earlier in the same article Keesing had demonstrated that except in some parts of eastern Oceania, particularly but not exclusively Eastern Polynesia, *mana* had little or no explicit connection with supernatural beings. Consequently he, like Pawley, shows an eastern bias in making these links. To Lichtenberk's statement about the absence of linguistic evidence for hereditary leadership it is necessary to add another caution: we have no linguistic evidence that POC society had a concept called *mana* that pertained either to gods and spirits, or to primogeniture. Neither do we know that life was thought to depend on "gods and ancestors". Though it might be argued that in the west concepts were affected by contact with NAN speakers, in fact gods and ancestors were considerably more important in some societies with NAN languages than among OC speakers (see Lawrence and Meggitt 1965; Chowning 1986).

### 2.3.2 KINSHIP

In discussing Blust's (1980) paper, Grace points out that kinship terms offer particular problems of reconstruction because the AN-speaking area contains so many different kinship systems. "As a consequence, the original meanings of some of the original terms must not have been preserved under any name in many Austronesian languages" (Grace 1980:235). This warning needs to be kept in mind in reading what follows. Some of the terms to be discussed (such as *\*tama* and *\*natu*) so consistently include the same kin in a wide variety of languages and within differing systems that there is no serious question about their original meaning in POC, though the range of meaning may be disputed. For some others, the suggestions are necessarily more tentative.

The first person to consider POC kinship was Milke, who in 1938 and 1958 reconstructed a series of terms based on those recorded for all of Oceania. Later he added a few terms that he thought were confined to the western end of the region, his Proto New Guinea Austronesian. Subsequent writings on the topic have tended to deal either with wider considerations, such as 'Early Austronesian social organization' (Blust 1980) or narrower ones, such as 'Sibling classification in Oceania' (Marshall 1984), rather than with complete systems. For POC, Milke's reconstructions are sometimes simply cited without much comment or alteration (e.g. Pawley and Green 1984:132). Yet it has long been clear that there are considerable differences between the kinship terminology of most of Melanesia and that of Polynesia. Pawley is the only person known to me who has discussed some of these differences, including pointing out how PPN diverges from what he calls both POC and PMN (1981:284). His chart nevertheless contains omissions, such as *\*rawa* (see below). It therefore seems worth discussing the Melanesian terms in more detail.

#### 2.3.2.1 CONSANGUINEAL TERMS

The consanguineal terms that I propose are as follows:

- (a) *\*t(i,u)mpu* – kin two generations removed (grandparent, grandchild). Although Blust (1980: 214) mentions the varying vowels in the PAN and PWMP root, I am not aware that the POC form has been so reconstructed, but there are a number of languages that contain *i*, ranging from some in Ross's North New Guinea cluster, such as Gitua *timbu-*, to Maori *tipuna*. It is necessary to propose an alternative to the usual *\*tumpu*. Although an alternative term for



'grandchild' has been proposed (see below), reflexes of the term just proposed are so often self-reciprocal that it is difficult to avoid attributing this meaning to POC, and indeed Blust (1980:214) attributes it to PAN.

- (b) *\*tama* – Fa, FaBr.
- (c) *\*tina* – Mo, MoSi.
- (d) *\*matuqa* – MoBr. Here I am accepting Milke's word for the distribution of this term, with this meaning, throughout Melanesia, from the New Hebrides and the Solomons to New Ireland, including Tolai of New Britain.
- (e) *\*(qa)lawa*. This term, with the shape *\*(a)lawa*, was originally reconstructed by Milke for 'SiCh, man speaking'. Later, because of Lakalai *hala*, he emended the reconstruction but kept the definition. But in Lakalai and other languages known to me that contain a reflex of the term, such as Kove *waha*, it is a self-reciprocal for MoBr as well as SiCh (man speaking). In Kove it is also the term for FaSi. I suspect that Milke was avoiding the problem of reconstructing two terms for the same kin type by so limiting his definition. With our present evidence, the double set seems unavoidable. We should, however, note that *\*matuqa* almost surely derives from the same root that means 'mature' (in Tolai, however, Lanyon-Orgill's dictionary shows a distinction between the adjective *matuka* and the kinship term *matua*-). It seems clear that in AN languages in general, *\*matuqa* referred to a wide range of senior kin (see Blust 1980; Chowning 1980) – as witness its use in parts of Polynesia as the general word for 'parent' (PPN *\*matuqa*). It may be that the MoBr, rather than some wider, referent for this should not be ascribed to POC. On the other hand, reflexes of the *\*qalawa* form seem rarely to mean anything else except MoBr or other cross-relative of the first ascending or descending generation. What seem to be almost wholly lacking are languages in which the term for MoBr reflects *\*matuqa* and that for SiCh reflects *\*qalawa*. (In some cases, especially with some Vanuatu languages, I am not sure whether either term reflects the proto-form.) It should be noted, however, that some of Milke's examples are misleading. Tolai *matua*, for example, is defined as including 'nephew or niece, also uncle or aunt' (Lanyon-Orgill 1960). Whatever the term may be, it is common to find that the one for MoBr is the same as that for SiCh. Examples in Rivers (1914) come from the Torres Islands, Ulawa and Lau, and various other parts of the Solomons. It is the case in Molima where the term, *ova*, is not cognate, as well as in Lakalai and Kove, and seems on the whole (judging from my recollection of collecting kinship charts at the University of Papua New Guinea) to be a common pattern in western Melanesia. If it is historically correct that *\*matuqa* designated only cross-kin of senior generation and *\*qalawa* only those of junior generation, then the widespread development of self-referential terms in the west led to either of these applying to both generations (see Torres Islands *meru/maru* as an example which presumably reflects *\*matuqa*). In this respect languages from Vanuatu eastwards seem to contrast with those spoken further west.<sup>13</sup>
- (f) *\*natu* – child of self and siblings, especially those of one's sex. Milke and Pawley have both discussed this as a specifically Melanesian term, although Blust has suggested that its use extends westwards. This is a term that does not appear in Polynesia.

<sup>13</sup>In addition to these, Blust postulated POC *\*dawa* (now *\*rawa* pers. comm.) 'MoBr/SiCh (man speaking) (recipr.)' together with a "Proto-Eastern Malayo-Polynesian term *\*dawa* ... meaning 'child-in-law'". These data are cited as part of his argument that at the PMP level, a single term was used for SiSo and DaHu (man speaking) (Blust 1980:213-214). Because I think, unlike Blust, that *rawa* was primarily an affinal term, I have discussed it under that heading.

- (g) *\*makumpu* is listed by Pawley as a second term for 'grandchild' together with his *\*tumpu*. The former term is almost wholly limited to regions from Vanuatu eastwards; cf. PPN *\*makupu/na*. Milke, however, records it for the Admiralties as well, but that still leaves it outside WOC as defined by Ross (1988). Nevertheless, if Admiralties languages developed quite separately from those spoken further east, as Ross assumes, then any form found both in the Admiralties and Polynesia must be POC. Furthermore, Blust (pers. comm.) tells me that there "are also non-OC reflexes of PMP *\*makempu* 'grandchild'". Here again we find a tendency in the western part of Melanesia, including all of the Solomons (see examples in Rivers) towards a single self-referential grandparent/grandchild term (or occasionally two terms reflecting sex differences), as opposed to a pattern in which the generations are distinguished. It is only a tendency, in that separate terms are used in many areas, but I know of no examples in which a *\*makumpu* term is used for 'grandparent'.
- (h) *\*t(o,u)ka* – older sibling, especially of same sex. This was originally reconstructed by Milke as *\*tuqaka*. Blust rejected this and proposed *\*toka* instead, together with *\*tua*, for POC (1984a:626). He does not discuss the similarities between these, but in (1984c:202) gives *\*tuka* alone. A number of Solomons languages justify reconstruction of a form in which *-u* is followed by a stop, so for the present this seems the best solution.
- (i) *\*tansi* – younger sibling, especially of same sex. This term has been accepted from the time of Milke's original formulation on.
- (j) *\*maRuqane* – brother, woman speaking (Blust's spelling; Pawley has *\*mwaqane*).
- (k) *\*(pa)pine* – sister, man speaking. I have modified Pawley's spelling, which lacks brackets, to allow for those languages such as Lakalai (*hata-vile*) in which the first syllable is absent, as is the case in most Polynesian languages (see list in Marshall 1984). I had earlier argued (Chowning 1984) that because these terms literally mean 'male' and 'female', their application to cross-sex siblings might have represented multiple innovations in different Oceanic languages, but I have been persuaded by others commenting on Marshall (1984), and by the distribution of these terms, that both (j) and (k) should be attributed to POC.
- (l) *\*l(i,o)pu* – cross-sex sibling. Milke originally reconstructed this, as *\*libu*, *\*lobu*, for PNGA (1965), and essentially he was right. With one possible exception, Kosrae of Micronesia, where the term is (*ma*) *louk* (Marshall 1984:621), terms reflecting this are confined to the north coast of New Guinea, neighbouring parts of New Britain, and offshore islands, including those of Milne Bay. Vitu *livu*, cited by Milke (1965:345), is the only example very far from New Guinea. Ross (1988:187) has suggested that *\*lopu* is the POC form and *\*lipu* an innovation in his Proto North New Guinea. Since he assigns Vitu to Meso Melanesian, I assume that he is postulating borrowing in that case, as I am also inclined to do. Unless the Kosrae example can be demonstrated to be cognate, I am reluctant to assign this word to POC. At best, it belongs to western Melanesian.

It has been impossible to reconstruct a proto-form for 'cross-cousin', even though many languages of Melanesia (as opposed to Polynesia) distinguish them terminologically from siblings and parallel cousins, who are almost always (except in Crow and Omaha systems) called by the same terms as siblings.

Pawley, following Blust, has also included, though with a query, a term for FaSi *\*aya*. I have already explained (Chowning 1980:222), my objections to this reconstruction, and so I have not included it in my list.

## 2.3.2.2 VOCATIVE TERMS

All of the terms that I shall be discussing have already been reconstructed by others. All I am doing here is bringing the material together, and in one or two cases arguing for the exclusively vocative nature, or the best gloss, for the term. These are the words that anthropologists call 'address terms', used in speaking directly to the person, as opposed to 'reference terms'. In my experience of looking at word lists for languages that I know, the former are often collected, doubtless because the questioner asks, 'What do you call your father?' and receives the equivalent of 'Daddy'. Within Melanesia, address terms vary more than do reference terms, and carelessness in this matter obscures the presence in many languages of reference terms derived from POC. Nevertheless, it is possible to reconstruct several terms for POC.

- (a) *(m)pu(m)pu* 'grandparent, grandchild'. This appears in Grace (1969) with a credit to Capell (1943), as *\*(mpu)mpu*. More recently Blust has provided evidence that the term *\*bubu* (of interest to him because it unites kin of 'alternating generations' as the reference terms he reconstructs do not) can be reconstructed for PAN (Blust 1980:219). This term extends throughout Melanesia, and also occurs in Polynesia (e.g. Tikopia *pu*). My modification of the proto-form reflects the fact that in much of western Melanesia, the vocative form does not seem to reflect the presence of the nasal, even when the reference term does. See for example Lakalai *pupu* (voc.), *tubu-* (ref.); Gitua *pupu* (voc.), *timbu-* (ref.); Molima *pupu* (voc.), *tubu-* (ref.); Namatanai *pupu* (voc.), *tubu-* (ref.) (Peekel cited in Rivers 1914). The exception is the one case cited by Capell, that of Motu *bubu e* (voc.), *tubu-* (ref.). In eastern Melanesia, as in Polynesia, neither term reflects the presence of a nasal in the proto-form, with the possible exception of Fiji, where a term *mbu* is used to address the grandmother only.<sup>14</sup> (Here I have ignored the fact that in a few cases, including Motu, there is no evidence in the publications that the term is also used for descending generations. It was a long time before I heard that usage in Lakalai, and I suspect that some of the data cited by others may not be complete.) Were it not for the discrepancies between vocative and reference forms in so many western Melanesian languages, and for the non-Oceanic witnesses, I would be inclined to assume that this term represented an abbreviation of the reference term (see below). Blust (1980:214) argues that for PWMP, the vocative term is the original, the *\*t-* of the reference term being an accretion, so reversing my argument. His vocative contains a nasal.
- (b) *mama* 'father'. This appears in Grace-Lincoln, glossed as 'FaBr, Fa, daddy', with a reference to a PAN term *\*mama* 'MoBr'. Witnesses include Gedaged, Motu, and Roviana. Even a cursory search of the literature indicates that the gloss should indeed be only 'daddy'; *mama* is the vocative term for 'father' along the entire Rai Coast, as in Gitua, and also in Lau and in Ngela. For possible origin, see under (c). Blust (1981) attributes this form to PMP.
- (c) *nana* 'mother' was reconstructed by Blust as a possibility, but he also presents the evidence for a competing form *\*ña[ ]* which can more certainly be given "the meaning 'mother' (voc.)" (Blust 1978:57-58). Together with *\*mama* and, more tenuously, with *\*(m)pu(m)pu*, this suggests a pattern in which a term – which might also have been a baby-talk form – was composed by reduplication of the second syllable of the reference term.

<sup>14</sup>Blust (pers. comm.) has reminded me of Fijian *i tubutubu* 'origin, parents, ancestors'. I had been assuming that the term derived from the verb *tubu* (from POC *\*tupu* 'grow'), under which it is listed, but clearly it may well derive from *\*(i,u)mpu*.

- (d) In addition to these, Milke (1965) mentioned a special vocative term, *\*wawa*, to accompany the MoBr *\*waya* reconstructed for PNGA. I have argued that *\*waya* may be related to his POC *\*qa(lawa)* (Chowning 1973), but as far as I know this vocative is indeed restricted to the general area for which Milke recorded examples – the north coast of New Guinea, extending over to New Britain (Kove *waha-* (ref.), *wawa* (voc.)). Whatever the status of the reference term, the vocative term seems to coincide roughly with Ross's North New Guinea cluster, and may be an innovation of that group.
- (e) Milke (1958) also reconstructed *\*kaka* as the address term for 'older sibling'. Since Blust, in rejecting Milke's reconstruction of the reference term *\*tuqa(ka)* has proposed another, *\*toka*, that ends with the same syllable (Blust 1984a:626), the argument that this fits the pattern of reduplicating the final syllable might still be made. The vocative term is widespread in Melanesia, at least from Motu to Bugotu, and because Blust (1984a:626) records its presence outside the OC languages, there need be no question about attributing *\*kaka* to POC.

### 2.3.2.3 AFFINAL TERMS

Only four of these have been reconstructed, and three appear as Melanesian terms in Pawley and Green. The only point of interest here is the distribution of two of the terms.

- (a) *\*ansawa* – 'spouse'. This was reconstructed by Milke and is fully acceptable.
- (b) *\*qipar* – 'spouse's sibling of opposite sex' (man's brother-in-law, wife's sister-in-law). Milke had reconstructed this without the initial consonant, but it is required by a considerable body of data (e.g. Lakalai *hiva*, admittedly with a doublet *iva*). Grace-Lincoln, though they omit the initial consonant, note that Dempwolff reconstructed it. Here I follow the spelling in Ross (1988), but note that there is some evidence for *\*q(e,i)paR*. See, for example, Kove and Molima *eya*. (Alternatively, this may reflect a different western term.)
- (c) *\*puŋao* – 'parent-in-law'. Here I follow the spelling in Pawley (1981) and Pawley and Green (1984); Grace-Lincoln have *\*puŋo*. There is no difficulty with the meaning of this term, originally reconstructed by Milke, but with its distribution, which seems to be wholly confined to eastern Melanesia and Polynesia. Competing with it in the west is:
- (d) *\*rawa* (Ross's spelling, with the same meaning). Possibly it should be *\*raua*, to account for Lakalai and Shortlands *loa* and Kove *lauwa*. Milke, who first proposed the term, thought that it was restricted to his PNGA (1965:345). I pointed out that Rivers had recorded it well away from New Guinea (Chowning 1973; see also Blust 1980:213) but I had not appreciated that reflexes of *\*rawa* seem to be wholly confined to the western region as delimited by Ross and by Tryon and Hackman. That is, they include the northern Solomons but do not extend south and east. On present evidence, *\*rawa* and *\*puŋao* are in complementary distribution.

### 2.3.2.4 DESCENT GROUPS

There is no agreement at present about whether descent groups can be ascribed to PAN society, and if so, what they were like (see Blust 1980, including Comments, and also Fox 1988). Nevertheless, they are almost universal in Melanesia, as well as in Polynesia and Micronesia (for a survey, see Oliver 1988). Blust has suggested that PAN *\*Rumaq* 'house' had a secondary meaning 'lineage', but gives no examples from Melanesia. I do not know of any, and have the impression that

the most common metaphorical extension is 'rope', in NAN as well as AN languages. Two terms relating to kinship groups have been reconstructed for eastern Oceania only. One is *\*qapusa*, defined in Grace-Lincoln as 'to sire a family, accumulation, descendants'. The second gloss seems intended (by Milke, who proposed this in 1961) to justify connections with PAN *\*qabusa*, on the one hand, and with Arosi *qahuta* 'all, whole' on the other. Neither of these has anything to do with kinship, though they (if indeed they are cognate with each other) certainly have cognates elsewhere in Melanesia, such as Molima *yawu-* 'all' (with a suffix varying for number). The comparative data point to a very narrow distribution for reflexes of *\*qapusa*, which only in Fiji, *yavusa*, refers to a descent group. I have, however, postulated that Kove *kabu* 'patrilineage' may be cognate. The shape of the Kove word, which is the same in Kaliai (Lusi), does not suggest direct inheritance, but I do not know which other language in West New Britain might have been the source. If the Kove and Fiji terms are indeed cognate, then they would justify a gloss referring to descent groups for POC.

Goodenough (1955) pointed to cognate terms in Micronesia and Polynesia describing descent groups, and Pawley has recently written more on the subject. He reconstructs the term as *\*kainaga* or *\*mata kainaga* and defines it as a 'landholding descent group'. He suggests that the term is particularly likely to be derived from POC *\*kai(n)* 'native of a place', attested with that meaning for Fiji and Arosi – and so of course raises doubts about whether it really refers to a descent group. Certainly it does not in many Polynesian languages (see list in Pawley 1981), though where it does not, it can, as he notes, label those "under the authority of a chief". While he says that "we may reconstruct *\*kainaga* 'high-order descent group' in the language immediately ancestral to Polynesian and Nuclear Micronesian" (1980:237), he does not mention Efate *kainaga*. Macdonald (1907) had proposed that the term was cognate with Samoan *'aiŋa*. If this is not a borrowing from a Polynesian language, then it certainly constitutes a Melanesian witness. It is, however, only because Pawley (1980:237) believes that "Polynesian and Nuclear Micronesian branches diverged perhaps 4,000 years ago" that he seems willing to attribute this term to POC, with the implication that the term dropped out throughout Melanesia, apparently because the hereditary chieftainship which he also attributes to POC society dropped out as well. But since hereditary chieftainship is in fact found in various parts of Melanesia that lack a term cognate with *\*kainaga* (Mekeo, Manam, the Trobriands, Kove, just to mention a few examples on or near New Guinea), this scenario is not persuasive. For the present, I prefer to attribute *\*kainaga* to a lower level eastern branch than to POC. It is worth noting that Goodenough has more recently (1985) suggested that the term originally referred to a land-holding unit rather than a descent unit.

It is just possible that one other term points to the existence of descent groups. POC *\*tuRaŋ* is defined in Grace-Lincoln as 'companion', and this is certainly its meaning in a number of languages (Motu, Wedau, Gedaged, Roviana, Nggela). In some languages, however, it is a kinship term. Examples include Proto Malaita *\*ula-* 'kin term – sibling, distant classificatory kinsman' (Levy and Smith 1969); Tolai (with matrilineal descent) *tura* 'man's brother, MoSiSo' (apparently). In Lakalai *tula* includes 'co-wife', women whose husbands are brothers, men whose wives are sisters (though Lakalai lacks moieties). In Bola it is the term for cross-cousin. The term might originally have referred to a member of one's descent group, but also perhaps to collateral kin of any sort. Apart from indicating that in POC society, there was a special relation between MoBr and SiCh (man speaking) – a situation found in a variety of kinship systems, including non-unilinear ones – all that can be clearly deduced from POC kin terms is that probably POC society did not prescribe marriage with kin in certain categories; otherwise the terms for affines would be the same as terms for 'consanguineal' kin. (Here I stick to the position in Chowning 1980, contra Blust 1980, that the POC terms for MoBr, etc. and child-in-law, etc. are unrelated.)



Even if we could attribute *\*kainaga* or *\*qapusa* to POC, we would have no clue as to what kind of descent group or category might be represented. For the former, cognate terms designate a patrilineal group in Tikopia, a matrilineal one in Truk, and a cognatic one in Maori. Furthermore, the kinship terms reconstructed so far for POC do not solve the problem. If POC society had unilinear descent, one would expect either that a term for cross-cousin was reconstructible, or that many more of the societies would be like Truk and the Trobriands in having kinship systems (Crow, in these cases) that group cross-cousins with other kin types. I would also expect a reconstructible term for FaSi unless, as in Kove, she was called by the same term as MoBr, but this does not seem to happen in many Melanesian societies. Citing Indonesian evidence, Fox (1988:40) has pointed out that "it is possible to construct either a two-line terminology or a cognatic [bilateral, without descent groups] terminology from much the same array of terms". I endorse his warning about the difficulty of drawing inferences about wider aspects of kinship organisation from simply examining the terms.

### 2.3.3 GENDER RELATIONS

Another Melanesian characteristic is so-called 'gender antagonism'. Although it is common elsewhere in the Pacific for women to suffer certain restrictions – for example, to have to avoid contact with canoes used for deep-sea fishing – the belief that contact with women endangers not only a wide variety of masculine activities but also male health is particularly strongly developed in Melanesia. It does not occur in all societies, but is just as prominent where languages are OC as where they are not. The strongest restrictions are associated with menstruation and childbirth, particularly the blood shed in childbirth. Consequently I was interested to discover that the words for menstruation in Kove and Molima were reduplicated forms of *\*tampu*: Kove *tavutavu*, Molima *taputapu*. I had expected to find similar terms in other languages, but have not succeeded. Possible though it might be that this meaning was a natural development from the term *\*tampu* (see section 2.2), it must also be noted that in Malaita, a similar term has precisely the opposite meaning. For example, Kwaio *abuabu* designates the 'sacred area beside men's house' (which must be kept free of female pollution) (Keesing 1975, 1982). It is consequently impossible at present to reconstruct a PMN term *\*tamputampu* with a unified meaning: presumably these are separate innovations in eastern and western Melanesia.

### 3. CONCLUSIONS

This exercise has produced conclusions that I had not anticipated. The most important is that the lexical data largely support Ross's (1988) hypothesis, based wholly on non-lexical data, of a division between what he calls Western Oceanic and other languages, lying mostly to the east, which he believes represent early departures from the New Guinea region. He also excludes from WOC the languages of the Admiralties and Mussau-Emira. These were classed with those of Polynesia and Micronesia by Milke (1958), but though Ross thinks that there may be a case for Micronesian connections he agrees with Blust (1978) that the languages of the Admiralties form a first-order subgroup of OC.

Because of the difficulty of distinguishing innovations that spread through the western region from retentions from an ancestral language, I am willing to label terms attested only for the west WOC only if it is understood that they may never have existed in the ancestor to the eastern languages and to those of the Admiralties. My data on the Admiralties are too poor to enable me to judge the position of those languages. As I noted in the Introduction, if for WOC (Ross's PWO) forms cognates are



found in languages to the west of New Guinea, then there is no difficulty about labelling terms POC even if they are no longer attested in the east. I also expect that cognates will be found in the east for many of the terms that now seem to be confined to western Melanesia, and such evidence would also warrant changing the label to POC. We need, however, to remember Ross's point about borrowing within the Solomons on the border between his two major groups; the evidence of cognates should come from further south.

The definition of EO as proposed initially by Biggs (1965), modified by Pawley (1972), and later amended by Pawley himself and other linguists, has led to uncertainty about its boundaries that leave me reluctant to use the term for the whole OC-speaking region exclusive of Ross's WOC. As modified by Lynch and Tryon (1985), 'Central Oceanic' includes all those OC languages excluded by Ross from WOC except those of the Admiralties. Whatever decision is made about terminology, it follows that PEO must not be equated with POC, and that the question of subgroupings at levels below POC or even PEO, must be kept firmly in mind, so that reconstructions can be labelled in accord with the present evidence. A system of numbers like that which Blust uses for PAN and its subdivisions would be a useful step in the right direction.

I should add that I would no longer argue, as I have in the past (e.g. Chowning 1973) that Lakalai, much less Kove, has particularly close ties with any EO language such as those of the south-east Solomons and Vanuatu. Accordingly, I feel that I can justifiably attribute to POC any pair of cognates shared between Lakalai or Kove and one of these eastern languages (see below).

Accepting Ross's WOC grouping solves certain problems of distribution that long puzzled me: sets of cognates widespread in languages that are not closely related to each other but are spoken in the general New Guinea region, and that seem not to occur elsewhere. An example is a word for Eclectus parrot: Molima *kanavala*, Kove *kahānani*, Sengseng *kahalaŋ*, Tolai *kalānara*; see also Kilivila *karaga*. Leaving aside some uncertainty about the order of the syllables, these are clearly cognate, but although the parrot itself is found in the Solomons, no similar words are recorded outside WOC, apart from a shorter term, *\*kara*, reconstructed by Blust for PEMP, and presumably reflected in Roviana *kara*. (I assume that POC *\*kaka* is not a much-shortened version.) Presumably the longer term postdates the departure of the ancestors of speakers of eastern languages. My data generally support the suggestion of Pawley and Green (1984:144-145) that a single dialect chain was long maintained in western Melanesia, but it requires that those who settled eastern Melanesia (and Polynesia and Micronesia) acquired new lexicon that differed considerably from some of that used in the west.

Lest it be thought that I am postulating a radical break between eastern and western Melanesia, I should add that my data on non-cultural lexicon show a considerable number of shared cognates which justify assigning to PMN/POC various terms which have either not been reconstructed yet (to my knowledge) or which have been assigned to smaller divisions such as PEO. Examples include the following: *\*ipu* 'head hair' (Lakalai *ivu*, Proto Malay *\*ifu*); *\*maqono(ta)* 'sweat' (Lakalai *maholo*, Raga *mamaono*, Navenevene and Tam (of Malekula) *mamonota*); *\*tue* 'kind of shellfish' (Kove, Mengen, Nggela, Fiji); *\*pote* 'willie wagtail' (Lakalai, Sa'a *pote*); *\*poŋipoji* 'morning' (a

reduplicated form of the word for 'night'); \*kesa 'green, blue'.<sup>15</sup> Geraghty (1983:369) notes the relation of Fijian *bogibogi* to PPN \**poŋipoŋi*, but my research, in collaboration with John Lynch, indicates that this is the only term for 'morning' that can be reconstructed for POC. Kove *vojivoŋi* is an example from the west. Geraghty who reconstructs PEO \*kesa 'dye', also cites as cognate Lakalai *kesa* 'painted with something unpleasant'. But that term means 'plastered with something unpleasant, such as excrement'; the cognates are the doublet, with reduplication, *kakesa*, *kakisa* 'green, blue'. Geraghty's gloss seems justified by the (non-Lakalai) cognates he cites, but the words mean specifically 'green, blue' in a number of widely separated languages, including Mota *gesagesaga*, and I think this is a justified gloss. I assume that unpublished data compiled by Blust, Pawley and others would add considerably to my own small collection of unpublished terms which are truly POC.

Finally, it needs to be remembered that here I am dealing with cultural items, assumed to be subject to borrowing much more than basic vocabulary is. We need not attribute a garamut-like term to the common ancestor of the separate 'clusters' that Ross (1988) calls North New Guinea and Meso Melanesian, which are spoken side-by-side, and indeed mixed in the Bola-Kove village of Kandoka. Wherever names for cultural items have a geographically limited distribution, regardless of whether the languages concerned belong to different first-order subgroups, we need to be very wary about assigning the item to an earlier stage of an ancestral culture. We know a good deal about the probable technology of POC society, but from linguistic evidence not a lot more, and in some cases less, than we could have deduced from the distribution of various items both in early historic times and in archaeological deposits. About other aspects of POC culture, we know very little. We do not know that the society possessed hereditary classes, chiefs or expert craftsmen, that special attention was given to firstborn children or seniority of descent lines, that there were gods or a concept of supernatural power called mana. We know that the people chewed betel, but it is much less certain that they drank kava. Apart from the special role probably given the MoBr, and the distinction of affines from consanguineal kin, we know no more about the kinship system than that they distinguished relative age and, almost surely, relative sex (compare Fox 1988 on Indonesia).<sup>16</sup>

If we are to understand POC culture, we must not apply that label to items that, even if supported by a large cognate set, are not attested outside the Central Oceanic region, or some division thereof. We need to keep in mind Grace's cautions about his original list, particularly his statement about the perceived advantages "in including rather than excluding sets" for which the evidence was less than ideal, and his use of keys to the "putative proto-language". It was to be expected that some of the reconstructions would be dropped, and many emended. He is not to be blamed if some of us have assumed too readily that everything on the lists is undoubtedly POC. Without his contributions our present knowledge of the POC lexicon, and our ability to draw inferences from it, would be a fraction of what it now is.

<sup>15</sup>I do not dispute Blust's reconstruction of another term, POC \**kadawa*, with the same meaning. POC certainly possessed true synonyms, as evinced by the two words for 'yesterday', \**ananapi* and \**noRap*.

<sup>16</sup>I do not, however, accept Fox's suggestion that there is no justification for the "common division of Austronesian terminologies into consanguineal and affinal components" (Fox 1988:43).

## BIBLIOGRAPHY

- BARRAU, Jacques**, 1958, *Subsistence agriculture in Melanesia*. Bernice P. Bishop Museum Bulletin 219. Honolulu: Bishop Museum.
- BENDER, Byron W.**, n.d., Micronesian languages. Mimeo.
- BIGGS, Bruce**, 1965, Direct and indirect inheritance in Rotuman. *Lingua* 14:383-413.
- BIGGS, Bruce, D.S. WALSH and J. WAQA**, 1970, *Proto-Polynesian reconstructions with English to Proto-Polynesian finder list*. Working Papers in Linguistics, Department of Anthropology. University of Auckland.
- BLUST, Robert A.**, 1978, *The Proto-Oceanic palatals*. Polynesian Society Memoir 43. Wellington: The Polynesian Society.
- 1980, Early Austronesian social organization: the evidence of language. *Current Anthropology* 21:205-247, 415-419.
- 1981, Variation in retention rate among Austronesian languages. Paper presented at the Third International Conference on Austronesian Linguistics, Bali, Indonesia.
- 1984, Austronesian etymologies – II. *Oceanic Linguistics* 22-23:29-149.
- 1984a, Comment on M. Marshall, Structural patterns of sibling classification in Oceania. *Current Anthropology* 25:626-628.
- 1984b, Malaita-Micronesian: an Eastern Oceanic subgroup? *Journal of the Polynesian Society* 93:99-140.
- 1984c, A Mussau vocabulary, with phonological notes. *PL*, A-69:159-208.
- 1986, Austronesian etymologies III. *Oceanic Linguistics* 25:1-123.
- BUCK, Peter (Te Rangi Hiroa)**, 1958, *The coming of the Maori*. Wellington: Whitcomb and Tombs.
- BULMER, S.**, 1982, Human ecology and cultural variation in prehistoric New Guinea. In J.L. Gressit, ed. *Biogeography and ecology of New Guinea*, 1:169-206. The Hague: Dr W. Junk.
- CAPELL, A.**, 1943, *The linguistic position of South-Eastern Papua*. Sydney: Australasian Medical Publishing Company.
- CHOWNING, Ann**, 1963, Proto-Melanesian plant names. In J. Barrau, ed. *Plants and the migrations of Pacific peoples* 39-44. Honolulu: Bishop Museum Press.
- 1973, Milke's "New Guinea cluster": the evidence from northwest New Britain. *Oceanic Linguistics* 12:189-243.
- 1977, *An introduction to the peoples and cultures of Melanesia*, 2nd edn. Menlo Park, CA: Chandler.
- 1979, Leadership in Melanesia. *Journal of Pacific History* 14/1-2:66-84.
- 1980, Comment on R.A. Blust, Early Austronesian social organization. *Current Anthropology* 21:184-185.
- 1984, Comment on M. Marshall, Structural patterns of sibling classification in Oceania. *Current Anthropology* 25:628.
- 1985, Rapid lexical change and aberrant Melanesian languages: Sengseng and its neighbours. In Andrew Pawley and Lois Carrington, eds *Austronesian linguistics at the 15th Pacific Science Congress*, 169-198. *PL*, C-88.
- 1986, Melanesian religions: an overview. *The encyclopedia of religion*, vol.9:349-359. New York: Macmillan.
- 1987, The supposed Austronesian content of the Trans-New Guinea phylum. In D.C. Laycock and W. Winter eds *A world of language: papers presented to Professor S.A. Wurm on his 65th birthday*, 119-125. *PL*, C-100.
- CHURCHILL, William**, 1911, *The Polynesian wanderings*. Washington: Carnegie Institution of Washington.
- 1916, *Sissano: movements of migration within and through Melanesia*. Washington: Carnegie Institution of Washington.
- CLARK, Ross**, 1979, Language. In J.D. Jennings, ed. *The prehistory of Polynesia*, 249-270. Canberra: Australian National University Press.
- CODRINGTON, R.H.**, 1885, *The Melanesian languages*. Oxford: Clarendon Press.
- DOUGLAS, Bronwyn**, 1979, Rank, power, authority: a reassessment of traditional leadership in South Pacific societies. *Journal of Pacific History* 14:2-27.
- FORTUNE, R.F.**, 1932, *Sorcerers of Dobu*. New York: E.P. Dutton.
- FOX, James J.**, 1988, Possible models of early Austronesian social organization. *Asian Perspectives* 26/1:35-44.

- FRENCH-WRIGHT, Renwick**, 1983, Proto-Oceanic horticultural practices. MA thesis, University of Auckland.
- GERAGHTY, Paul A.**, 1983, *The history of the Fijian languages*. *Oceanic Linguistics* Special Publication 19.
- GOODENOUGH, Ward H.**, 1955, A problem in Malayo-Polynesian social organization. *American Anthropologist* 57:71-83.
- 1985, Personal communication cited in Pawley 1985.
- GRACE, George W.**, 1969, A Proto-Oceanic finder list. *Working Papers in Linguistics, Department of Linguistics, University of Hawaii*, 2:39-84.
- 1976, History of research in Austronesian languages of the New Guinea area: general. In S.A. Wurm, ed. *New Guinea area languages and language study*, vol.2: *Austronesian languages*, 55-71. PL, C-39.
- 1980, Comment on R.A. Blust, Early Austronesian social organization. *Current Anthropology* 21:235.
- GRACE, George W. and Peter C. LINCOLN**, 1979, Proto-Oceanic cognate sets. Computer printout.
- GREEN, Roger C.**, 1979a, Early Lapita art from Polynesia and Island Melanesia: continuities in ceramic, barkcloth, and tattoo decorations. In S.M. Mead, ed. *Exploring the visual art of Oceania* 13-31. Honolulu: University of Hawaii Press.
- 1979b, Lapita. In J. Jennings, ed. *The prehistory of Polynesia* 27-60. Canberra: Australian National University Press.
- GROUBE, L.M.**, 1971, Tonga, Lapita pottery, and Polynesian origins. *Journal of the Polynesian Society* 80:278-316.
- HADDON, A.C. and J. HORNEILL**, 1975, *Canoes of Oceania*. Honolulu: Bishop Museum Press.
- HUDSON, Elizabeth**, 1989, All ngaun: Oceanic terms for 'dog'. In R. Harlow and R. Hooper, eds *VICAL 1: Oceanic Languages*, Part 1, 283-304. Auckland: Linguistic Society of New Zealand.
- KEESING, Roger M.**, 1975, *Kwaio dictionary*. PL, C-35.
- 1982, *Kwaio religion: the living and the dead in a Solomon Island society*. New York: Columbia University Press.
- 1984, Rethinking 'mana'. *Journal of Anthropological Research* 40:201-217.
- KIRCH, Patrick W.**, 1984, *The evolution of the Polynesian chiefdoms*. Cambridge: Cambridge University Press.
- KROLL, H.**, 1938, Sagen und Märchen der Bola. *Zeitschrift für Ethnologie* 70:371-425.
- LANYON-ORGILL, Peter**, 1960, *A dictionary of the Raluana language*. Victoria, B.C.: privately published.
- LAWRENCE, Peter and M.J. MEGGITT**, 1965, *Gods, ghosts and men in Melanesia*. Melbourne: Oxford University Press.
- LEENHARDT, Maurice**, 1946, *Langues et dialectes de l'Austro-Mélanésie*. Paris: Institut d'Ethnologie.
- LEVY, Richard and Nathan SMITH**, 1969, A Proto-Malaitan lexicon. Typescript.
- LICHTENBERK, Frantisek**, 1986, Leadership in Proto-Oceanic society: linguistic evidence. *Journal of the Polynesian Society* 95:341-356.
- LINCOLN, Peter C.**, n.d., On Gedaged, Milke, and Proto-Oceanic. Typescript.
- LYNCH, John and Darrell 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.
- MACDONALD, D.**, 1907, *The Oceanic languages: their grammatical structures, vocabulary, and origin*. London: Henry Frowde.
- MARSHALL, Mac**, 1984, Structural patterns of sibling classification in Oceania. *Current Anthropology* 25:597-637.
- MILKE, Wilhelm**, 1958, Ozeanische Verwandtschaftsnamen. *Zeitschrift für Ethnologie* 83:226-229.
- 1961, Beiträge zur ozeanischen Linguistik. *Zeitschrift für Ethnologie* 86:162-182.
- 1965, Comparative notes on the Austronesian languages of New Guinea. *Lingua* 14:330-348.
- NILES, Don**, 1983, Why are there no garamuts in Papua? *Bikmaus* 4/3:90-104.
- OLIVER, Douglas**, 1955, *A Solomon Island society*. Cambridge, Mass.: Harvard University Press.
- 1988, *Oceania: native cultures of Australia and the Pacific Islands*. Honolulu: University of Hawaii Press.
- PARKINSON, R.**, 1907, *Dreissig Jahre in der Südsee*. Stuttgart: Strecker und Schröder.

- PATON, W.F., 1973, *Ambrym (Lonwolwol) dictionary*. PL, C-21.
- PAWLEY, Andrew, 1972, On the internal relationships of Eastern Oceanic languages. In R. Green and M. Kelly, eds *Studies in Oceanic culture history*, vol.3, 1-142. Honolulu: Bishop Museum Press.
- 1978, The New Guinea Oceanic hypothesis. *Kivung* 21:99-151.
- 1980, Comment on R.A. Blust, Early Austronesian social organization. *Current Anthropology* 21:236-237.
- 1981, Melanesian diversity and Polynesian homogeneity: a unified explanation for language. In J. Hollyman and A. Pawley, eds, *Studies in Pacific languages and cultures in honour of Bruce Biggs*, 269-309. Auckland: Linguistic Society of New Zealand.
- 1982, Rubbishman commoner, Big-man chief? Evidence for hereditary chieftainship in Proto-Oceanic. In J. Siikala, ed. *Oceanic studies: essays in honour of Aarne A. Koskinen*, 33-52. Helsinki: Finnish Anthropological Society.
- 1985, Proto-Oceanic terms for 'person': a problem in semantic reconstruction. In V. Z. Acson and R. L. Leed, eds *For Gordon Fairbanks*, 92-104 *Oceanic Linguistics* Special Publication 20.
- PAWLEY, Andrew and Roger C. GREEN, 1984, The Proto-Oceanic language community. *Journal of Pacific History* 19:123-146.
- POLLOCK, Nancy J., 1984, Breadfruit fermentation practices in Oceania. *Journal de la Société des Océanistes* 79:151-164.
- RAY, Sidney H., 1907, *Linguistics. Reports of the Cambridge Anthropological Expedition to Torres Straits*, vol.3. Cambridge: Cambridge University Press.
- REICHARD, Gladys A., 1933, *Melanesian design: a study of style in wood and tortoise shell carving*. New York: Columbia University Press.
- RIVERS, W.H.R., 1914, *The history of Melanesian society*. 2 vols. Cambridge: Cambridge University Press.
- ROSS, Malcolm D., 1977, Relationships of the Austronesian languages of the Sepik and western Madang coast of New Guinea. Mimeographed.
- 1988, *Proto Oceanic and the Austronesian languages of western Melanesia*. PL, C-98.
- n.d., Lexicon of Proto Papuan Tip reconstructions. Computer print-out.
- SAHLINS, Marshall D., 1963, Poor man, rich man, big-man, chief: political types in Melanesia and Polynesia. *Comparative Studies in Society and History* 5:285-303. London.
- SCALETTA, Naomi, 1987, The first-born child as parental exemplar among the Kabana of West New Britain, Papua New Guinea. Paper delivered to the Association for Social Anthropology in Oceanic. Typescript.
- SELIGMAN, C.G., 1910, *The Melanesians of British New Guinea*. Cambridge: Cambridge University Press.
- SPRIGGS, Matthew, 1985, The Lapita cultural complex: origins, distribution, contemporaries and successors. In R. Kirk and E. Szathmary, eds *Out of Asia: peopling the Americas and the Pacific*, 185-206. Canberra: Journal of Pacific History.
- TERRELL, John, 1986, *Prehistory in the Pacific Islands*. Cambridge: Cambridge University Press.
- THURSTON, William R., 1987, *Process of change in the languages of north-western New Britain*. PL, B-99.
- TRYON, D.T. and B.D. HACKMAN, 1983, *Solomon Islands languages: an internal classification*. PL, C-72.
- WURM, S.A. and B. WILSON, 1975, *English findex of reconstructions in Austronesian languages* (post-Brandstetter). PL, C-33.
- YEN, Douglas E., 1973, Agriculture in Anutan subsistence. In D.E. Yen and Janet Gordon, eds *Anuta: a Polynesian Outlier in the Solomons*, 113-149. Pacific Anthropological Records 21. Honolulu: B.P. Bishop Museum.



THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...

THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...

THE SEARCH

THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...

THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...

THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...

THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...

THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...

THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...

THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...

THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...

THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...

THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE... THE FUTURE OF THE FUTURE...



## THE SEARCH FOR SUBJECT IN SAMOAN

KENNETH WILLIAM COOK

### 1. INTRODUCTION

Interest in which nominal of a Polynesian clause should be considered the subject dates back at least to Biggs (1974), who argued that the absolutive (and not the ergative) of a transitive clause in East Futunan (a language related to Samoan) should be considered the subject. Biggs also suggested that the same analysis might hold for other Polynesian languages (such as Samoan) that exhibit ergative case marking. However, Chung (1976, 1978), Cook (1978), Ochs (1982) and Voigt (1986) have either argued or assumed that the ergative (and not the absolutive) of a Samoan transitive clause is the subject. Both of these approaches assume that Samoan clauses have subjects. Mosel (1987), in contrast, argues that Samoan does not have the category subject. In this paper I will show that the evidence that Mosel presents does not necessarily lead to that conclusion. In addition, I will present new evidence for subject in Samoan. Many of my claims concerning whether Samoan has the category subject will involve the subject properties listed in Keenan (1976).<sup>1</sup>

In this discussion, I will use Dixon's (1979) A and O to refer respectively to the agent and patient of a transitive verb and S to refer to the primary participant of an intransitive verb. A language exhibits an ergative case-marking pattern if it marks S's and O's with the same (absolutive) case and marks A's with a special (ergative) case. I will also use 'absolutive' to refer to S's and O's and 'ergative' to refer to A's. A language shows an accusative pattern if it marks S's and A's with the same (nominative) case and O's with another (accusative) case. The conflation of S's and A's corresponds to the traditional term 'subject' and O corresponds to 'direct object'.

The terms 'accusative' and 'ergative' also refer to these types of patterning in phenomena other than case marking. Languages may (and often do) exhibit ergativity with respect to some phenomena and accusativity with respect to others. English, for example, is dominantly accusative but shows ergativity in nominalisations ('the shooting of the hunter/deer' versus 'the shooting of the deer by the hunter'). In addition, languages are in no way restricted to these patterns. There are other patterns such as those found in the so-called 'active languages' in which rules refer to semantic roles. Notice

---

<sup>1</sup>The theoretical approach I take in this paper differs from that in Cook (1988). I maintain, however, the claim expressed there that word order plays more of a part in Samoan syntax than previously assumed. The purpose of this paper is to argue for the category subject in Samoan (pace Mosel) and not to discredit the interesting observations concerning Samoan tendencies in coding semantic content that Mosel makes in her paper in this volume.

that evidence for accusativity in a language is also evidence for the category subject since it is in the accusative pattern that subject (as the conflation of S and A) is manifested.

Mosel (1987:455) roughly defines subject as a "grammaticalized topic within the clause" and argues against subject in Samoan by showing that a phenomenon in Samoan either follows an ergative pattern or, at least, does not follow an accusative pattern. Mosel's evidence comes from facts concerning Samoan case marking, agreement in number, grammatical indispensability, word order, clitic placement, coreferential deletion across conjunctions, relativisation, left-dislocation and equi-NP deletion. Some of these phenomena do reveal ergativity, but the majority, I will argue, reveal accusativity and, therefore, provide evidence for the category subject (as the conflation of S and A). In support of this claim I will offer further evidence concerning raising, imperatives and the coding of causers.

Having shown that there is syntactic evidence for the category of subject defined as the conflation of S and A, I will show that there is also evidence from quantifier float, reflexivisation, verb agreement, and leftmost position for subjects that are either deep but not surface or surface but not deep. In particular I will argue that the A of a V(erb)AO clause is both deep and surface subject while that of a VOA clause is a deep but not a surface subject. I will then propose plausible analyses for VAO and VOA clauses.

## 2. CASE MARKING

Samoan exhibits an ergative case-marking pattern. Both S's and O's are either left unmarked or marked 'o (absolutive). A's are in the ergative case (marked *e*).<sup>2</sup>

- (1) a. 'Ua alu le tama 'i Sāmoa.  
 PERF go the boy DIR Samoa  
 The boy has gone to Samoa.
- b. Na fufulu e le tama le ta'avale.  
 PAST wash ERG the boy the car  
 The boy washed the car.

The main focus of this paper will be on the canonical intransitive and transitive clauses exemplified in (1a) and (1b). However, upon occasion, I will also need to refer to middle and *lavea* clauses. Middle clauses contain middle verbs, which are, for the most part, verbs of emotion, perception and communication. The agentlike participant of a middle verb is an absolutive and the patientlike participant is an oblique complement. In Cook (1988:72ff.) I have argued that these verbs are intransitive, and I will treat them as such here. *Lavea* verbs are intransitive verbs that take a patientive (rather than an agentive) S. If an agent is encoded in a *lavea* clause (i.e. a clause containing a *lavea*

<sup>2</sup>There is variability concerning the presence of the ergative case marker. As Ochs (1982:660) points out, it is "used more by men than by women, and is more frequent in speech to non-family members than in speech among household intimates". In this paper I am assuming a level of usage in which the marker would appear in all of its possible contexts. The abbreviations used in the example sentences are as follows: ABS – absolutive, ACC – accusative, ANAPH – anaphoric pronoun, CAUS – causative, COMP – complementiser, DIR – directional, ERG – ergative, IMP – imperfect, INF – infinitive, LOC – locative, NEG – negative, NOM – nominative, OBL – oblique, PERF – perfect, PL – plural, PRES – presentive, PRO – pronoun, PROG – progressive, TR – transitive.

verb), it is in the locative case and not the ergative.<sup>3</sup> Sentences (2a) and (2b) respectively exemplify middle and *lavea* clauses.

- (2) a. *E alofa le teine 'i lona tinā.*  
 IMP love the girl DIR her mother  
 The girl loves her mother.
- b. *Na lavea le teine i le tama.*  
 PAST hurt the girl LOC the boy  
 The girl was hurt by the boy.

### 3. AGREEMENT IN NUMBER

Mosel (1987:457) shows that if a verb agrees in number with a nominal, it is with the absolutive. This is illustrated in (3) with the plural intransitive verb *ō* 'go' and plural transitive verb *nunuti* 'crush'. I will henceforth refer to this phenomenon as 'traditional verb agreement' and later in this paper I will present another version of agreement, which I will call 'modern verb agreement'. As can be seen in (3a) and (3b), traditional verb agreement follows an ergative pattern.<sup>4</sup>

- (3) a. *'Ua ō tamaiti 'i Sāmoa.*  
 PERF go(PL) children DIR Samoa  
 The children have gone to Samoa.
- b. *'Ua nunuti e le tama fuāmoa.*  
 PERF crush(PL) ERG the boy egg  
 The boy has crushed the eggs. (Milner 1966:159)

### 4. GRAMMATICAL INDISPENSABILITY

Keenan (1976:313) proposes that indispensability is a property of basic subjects. This is to say that in a context-free environment, "non-subjects can often be eliminated from a sentence with the result being a complete sentence". Thus it is possible to dispense with the object in *John hunts lions*, but not with the subject. One can say *John hunts*, but not *\*Hunts lions*.

Mosel observes that the agents of both transitive and intransitive verbs are dispensable, whereas the patient of a transitive verb is not. In illustration of these observations, she offers the following passage about net making in which the agent of intransitive *alu* 'go', and transitive *fai* 'make' and *su'e* 'look for' is left unspecified. In contrast, the patient of *su'e* is specified.

- (4) *Ona fai 'o le 'upega. Ona alu lea*  
 then make ABS the net then go (then)  
 Then the net is made. Then (someone) goes

<sup>3</sup>Robert Blust (pers. comm.) has pointed out that it is common in Austronesian languages for agents to be marked with cognates of the marker *i* that appears in (2b). (Samoan is an Austronesian language.) I have chosen to gloss this marker 'locative' because that is its most basic value.

<sup>4</sup>The absence of a singular determiner (either *le* 'the' or *se* 'a' or one of their combined forms such as *la'u* 'my') is an indication of plurality.

'ua su'e mai se isi  
 PERF look=for hence a one  
 (he) looks for some one

na te fai-a le 'upega.  
 he IMP make-TR the net  
 to make the net. (Krämer 1902-03, 2/180)

If I understand Mosel's argument, it runs along these lines: since subjects are supposed to be indispensable, A's and agentive S's like those in (4) are not subjects because they are, in fact, dispensable.

I beg to disagree with Mosel's conclusion. First, it is not clear to me that indispensability is a subject property. Keenan himself (1976:313) states that "several ergative languages...appear to permit unspecified subject deletion", and he cites Tongan as an example. (Tongan is closely related to Samoan.) Notice also that Keenan's English examples are valid only because the verb *hunt* can be used both transitively and intransitively. If we were to take a strictly two-argument verb like *resemble*, \**John resembles* would not be much better than \**resembles Bill*. Furthermore, there are accusative languages which allow unspecified subjects even in contextually independent sentences. Consider (5) from Japanese:

- (5) *Ano toshi ni Long Beach de Grand Prix o shita.*  
 that year in Long Beach in Grand Prix ACC did  
 That year (they) held the Grand Prix in Long Beach.

Although indispensability may not clearly be a subject property, Mosel's observations are still interesting because they suggest that dispensability (rather than indispensability) runs along the lines of accusativity rather than ergativity. If absolutes (and only absolutes) were dispensable, then dispensability would argue for ergativity, but as we have seen, S's and A's are dispensable while O's are not. Therefore, dispensability is an accusative phenomenon in Samoan.

## 5. BASIC WORD ORDER

Mosel (1987:458) describes the word order of basic Samoan clauses as rigidly verb initial and she claims that the order of nominals is pragmatically determined: "the participant which the speaker wants to focus as being the most important part of the new information directly follows the verb". I agree with Mosel's observation that Samoan is verb initial; however, I do not agree with her claim that the immediate postverbal position is one of new information. In support of her claim, Mosel (1987:459) offers a sequence of sentences (of which (6) is one) from a story in which a boy brings *mafu* 'cubes' of a certain Samoan dish to a woman who is attracted to him. The woman asks the boy to put the food into her mouth, which he does. Apparently for Mosel, the fact that the boy (and not the woman) puts the food into her mouth is new information. I would say that Mosel is confusing new information with contrast. Since the boy has been introduced in the telling of the story just before this event takes place, the boy cannot be new information. The fact that the boy (and not the woman) puts the food into the woman's mouth is, however, contrastive.

- (6) *Ona tu'u atu lea e le tama 'o le mafu*  
 then put forth (then) ERG the boy ABS the cube  
 Then the boy put the cube

*i le gutu o le tama'ita'i.*

LOC the mouth of the lady

in the mouth of the lady. (Krämer 1902-03, 1/267)

I would claim, instead, that unless contrast is involved, the nominal that appears in immediate postverbal position is usually one of *old* information. This fits the cross-linguistic observation that old information generally precedes new. Given that pronouns generally represent old information, this claim is also supported by the fact exemplified in (7) and (8) that postverbal pronouns occur in immediate postverbal position.

- (7) a. 'O le ā le mea na fai e le tama?  
 PRES the what the thing PAST do ERG the boy  
 What did the boy do?
- b. Na opo e ia le teine.  
 PAST hug ERG he the girl  
 He hugged the girl.
- (8) a. 'O le ā le mea na tupu 'i le teine?  
 PRES the what the thing PAST happen DIR the girl  
 What happened to the girl?
- b. Na opo 'o ia e le tama.  
 PAST hug ABS she ERG the boy  
 She was hugged by the boy.

In (7b) and (8b) the A and O are respectively old information since they have been previously mentioned in (7a) and (8a). However, if neither the A nor the O is previously mentioned, all else being equal, the A will precede the O.<sup>5</sup>

- (9) a. 'O le ā le mea na tupu?  
 PRES the what the thing PAST happen  
 What happened?
- b. Na opo e le tama le teine.  
 PAST hug ERG the boy the girl  
 The boy hugged the girl.

This observation indicates that VAO (and not VOA) word order is basic. Here are some other reasons for believing that VAO is basic: consultants spontaneously translate basic English transitive clauses with VAO word order. When teachers in Western Samoa write model transitive sentences for their students, they regularly order the A before the O (Mosel, pers. comm.). In their textbook written to reinforce the Samoan language among American Samoan bilingual students, Tuitele, Sâpolu, and Kneubuhl (1978:41-50) state that both word orders are possible, but they show a strong preference for VAO word order in their examples and exercises. In his statistical analysis of the frequencies of word order types in everyday conversation and *talanoaga* (conversations that occur in *fono*, i.e. village meetings), Duranti (1981:section 8.2.3) found VAO word order to be the most

<sup>5</sup>Ochs (1982:664) observes that word order also varies with social context: when people talk to family members, they use VOA word order more frequently than VAO. The opposite is true when they speak to outsiders. Also, men use VAO word order more frequently than women. This is essentially the same distribution as that for the presence of the ergative marker mentioned in footnote 2. The social context that I am assuming in this paper is that of men speaking to outsiders.



common in clauses with all three constituents present. Chung and Milner have both expressed similar opinions. Chung (1978:14) claims that the basic word order of Samoan (as well as other Polynesian languages) is verb-subject-object-oblique, and Milner (1973:636) reports that the 'actor' tends to precede the 'goal' in what I am calling transitive clauses.

Intransitive clauses normally order their S's before complements (cf.(1a)). There are, of course, exceptions, among which are presentative clauses, which allow their S's to precede or follow complements.<sup>6</sup>

On the basis of these observations, I conclude that S's and A's precede other nominals in the basic word order of Samoan. The pattern involved here, then, is an accusative one in that S's and A's both occur first after the verb in the basic word order. Not only does the accusativity of this phenomenon argue for the category 'subject' in Samoan, but there is also the fact that S's and A's occur *leftmost* among the postverbal nominals. As Keenan (1976:319) observes, "basic subjects are normally the leftmost occurring NP in basic sentences".

## 6. CLITIC PLACEMENT

Clitic placement, as described by Chung (1978:34-37), moves a postverbal personal pronoun into a position between the tense-aspect marker and the verb. Clitic placement derives (10b) from (10a):<sup>7</sup>

- (10) a. 'Ua ō lātou 'i Sāmoa.  
PERF go(PL) they DIR Samoa  
They have gone to Samoa.
- b. 'Ua lātou ō 'i Sāmoa.  
PERF they go(PL) DIR Samoa  
They have gone to Samoa.

Sentence (10) shows that the S's of canonical intransitive clauses undergo clitic placement. Sentences (11) and (12) illustrate that A's but not O's also undergo this rule:

- (11) a. Na opo e ia le teine.  
PAST hug ERG he the girl  
He hugged the girl.
- b. Na ia opo-ina le teine.<sup>8</sup>  
PAST he hug-TR the girl  
He hugged the girl.
- (12) a. Na opo 'o ia e le tama.  
PAST hug ABS she ERG the boy  
She was hugged by the boy.

<sup>6</sup>See Mosel (1987:476) for more on presentative clauses. See also Cook (1988:79ff.) for some discussion of what I have called *galo* clauses. These clauses also allow verb-absolutive-oblique and verb-oblique-absolutive word order.

<sup>7</sup>If the tense-aspect marker is the imperfect marker *e*, the clitic and the tense marker are inverted and the tense marker becomes *te*. The preverbal pronouns involved here have been called 'clitics' and 'proclitic pronouns' probably because they appear in a particular position before the verb and because many of them are phonologically reduced versions of the corresponding postverbal pronouns. For example, 'ou and 'e are respectively the clitic versions of 'a'u 'I' and 'oe 'you'.

<sup>8</sup>When an ergative pronoun is moved into clitic position, it loses its case marker and generally a form of the so-called transitive suffix (either *-a* or *-ina*) is suffixed to the verb. See Cook (1988:155ff.) for more on this matter.

- b. \**Na ia opo(-ina) e le tama.*  
 PAST she hug-TR ERG the boy  
 (She was hugged by the boy.)

Notice also that the patientive S of a *lavea* clause cliticises while the agent of such a clause does not. Thus clitic placement affects S's and A's but not O's (nor agents that are not S's or A's).

- (13) a. *Na 'ou lavea i le tama.*  
 PAST I hurt LOC the boy  
 I was hurt by the boy.

- b. \**Na ia lavea le teine.*  
 PAST he hurt the girl  
 (He hurt the girl.)

Mosel (1987:460-461) refers to this phenomenon as "pronominalization by proclitic pronouns" and admits that it follows an accusative pattern. She points out, however, that there are constraints on this rule if the verb is intransitive. Some speakers do not allow third person singular pronouns to cliticise if the clause is intransitive (Chung 1978:221), and third person dual and plural pronouns generally do not cliticise if the verb expresses a state (Johnson and Harmon n.d.:210). In my own work on clitic placement I have found that S's do not cliticise if they are what Langacker (1987:34) calls 'zeros', that is, participants that do not move, have an experience, or undergo a change of state. Sentence (14) illustrates this last constraint. Since the S of *sa'o* '(be) right' is a zero, it cannot undergo clitic placement. Below we will see that this constraint on zeros also affects a certain version of equi-NP deletion.

- (14) a. *E sa'o 'oe.*  
 IMP right you  
 You are right.
- b. \**E te sa'o.*  
 you IMP right  
 (You are right.)

I do not find it problematic that this or any other accusatively organised rule has a semantic constraint. Dixon (1979:108), whose definition of subject we are working with here, sees the potentiality for being an "initiating/controlling" agent as the "semantic link" between S and A. He even asserts, "Ideally we should define 'subject' as {A, S<sub>a</sub>}, linking A with the subtype of S which can be agent." His reason for not defining subject as {A, S<sub>a</sub>} is that "no language consistently distinguishes S<sub>a</sub> from S<sub>o</sub> in all aspects of its grammar." S is generalised, according to Dixon, to the extent that it includes S<sub>o</sub>. As for Samoan clitic placement, the conflation of A's and agentive S's apparently is generalised to include patientive but not zero S's.

## 7. COREFERENTIAL DELETION ACROSS CONJUNCTIONS

In studies on ergativity, it has become common to ask which nominal (if either) is understood as the S associated with the second verb phrase in a sentence like *The man hit the woman and ran away*. If this phenomenon follows an accusative pattern, the man is understood as the person who ran away. If it follows an ergative pattern, then the woman is the person who ran away. In other words, in the accusative pattern the A associated with the first verb phrase is the controller of the deletion of the S. In the ergative pattern, the O is the controller. As Mosel (1987:462-464) shows, testing for

this phenomenon in Samoan reveals neither ergativity nor accusativity. The S in the second clause of (15) cannot be deleted without creating ambiguity.<sup>9</sup>

- (15) *Sā fasi e le tamāloa le fafine*  
 PAST beat ERG the man the woman  
 The man hit the woman
- ona sola ai lea.*  
 then run=away ANAPH that  
 and (?) ran away.

Example (15) is ambiguous in that the deleted S of the second clause could be coreferential with either the A or the O of the first clause. Notice that this fact does not necessarily mean that Samoan lacks the category subject. The same ambiguity could (and does) result in a language such as Japanese which clearly has subject. As illustrated in (16), conjoining clauses in Japanese with the conjunction *ga* 'and/but' results in the same type of ambiguity.

- (16) *Taroo ga Hanako o butta ga nigeta.*  
 Taroo NOM Hanako ACC hit and/but ran=away  
 Taroo hit Hanako and/but (?) ran away.

## 8. RELATIVISATION

There are two general strategies of relativisation in Samoan: one involves pronominalisation and the other, deletion.<sup>10</sup> When obliques and possessives are relativised, they follow the pronominalisation strategy in that they usually leave behind a pronominal copy.

- (17) a. *le aso na sau ai le tama*  
 the day PAST come PRO the boy  
 the day the boy came
- b. *le tama 'ua pē lana maile*  
 the boy PERF die his dog  
 the boy whose dog has died (lit. the boy his dog has died)

When S's and O's are relativised, they follow the deletion strategy in that do not leave behind a pronominal copy.

- (18) a. *le tama na sau ananafi*  
 the boy PAST came yesterday  
 the boy who came yesterday

<sup>9</sup>The normal way of conjoining clauses which refer to sequential events is with the expression *ona* [verb] *ai lea* '(and) then', which appears in (15). According to Churchward (1951:27), the conjunction *ma* 'and' "joins words rather than phrases". Note also that the presence of a third person singular pronoun in the second clause of (15) does not disambiguate the clause.

<sup>10</sup>As Mosel (1987:464) observes, there are also two types of relative clauses in Samoan: syndetic and asyndetic. The former employs the relative pronouns *le* (singular) and *ē* (plural), and the latter does not. The two do not differ structurally or semantically when used attributively. Hence Mosel limits her discussion to asyndetic relative clauses and I will do the same here.

- b. *le mea na fai e le tama*  
 the thing PAST do ERG the boy  
 the thing that the boy did

A's optionally leave behind a clitic pronoun.

- (19) *le tama na (ia) opo-ina le teine*  
 the boy PAST he hug-TR the girl  
 the boy who hugged the girl

For Mosel, the A clitic would have to be obligatory in order for this pattern to be ergative. If I understand her reasoning, she feels this way because in the cases in which the ergative clitic fails to occur, A's behave like S's and O's in that they do not leave behind a pronominal copy. Given that S's and O's behave exactly alike in that they follow exactly the same (deletion) strategy (while A's follow an optional pronominalisation strategy), it is my opinion that with respect to relativisation, Samoan shows ergativity.

## 9. LEFT DISLOCATION

Left dislocation is a syntactic process in which a nominal is extracted from the clause and moved to the front of the sentence. Once moved, the nominal is normally preceded by 'o, which Churchward (1951:21) calls the presentive particle. One of the functions of left dislocation is that of singling out an individual or a point of time, etc.

Left dislocation is similar to relativisation but the breakdown as to what nominal type follows which strategy is slightly different. Obliques and possessives normally leave behind a pronominal copy.

- (20) a. *'O le aso na sau ai le tama.*  
 PRES the day PAST come PRO the boy  
 (It is) the day the boy came.
- b. *'O le tama 'ua pē lana maile.*  
 PRES the boy PERF die his dog  
 (He is) the boy whose dog has died. (lit. the boy his dog has died)

O's follow the deletion strategy.

- (21) *'O le mea na fai e le tama.*  
 PRES the thing PAST do ERG the boy  
 (It is) the thing that the boy did.

S's and A's optionally leave behind a copy in the form of a clitic pronoun.<sup>11</sup>

- (22) a. *'O 'oe na ('e) sau ananafi.*  
 PRES you PAST you come yesterday  
 (It is) you who came yesterday.

<sup>11</sup>Third person singular S's do not leave behind a pronoun. This corresponds to the fact that some speakers disallow third person clitics in intransitive clauses (Mosel 1987:466).

- b. 'O le tama na (ia) opo-ina le teine.  
 PRES the boy PAST he hug-TR the girl  
 (It is) the boy who hugged the girl.

Since S's and A's behave similarly in that they optionally leave behind a pronominal copy (while O's follow the deletion strategy), left dislocation exhibits an accusative pattern.

Mosel claims that left dislocation is not accusatively organised because left dislocation of O's represents the unmarked case and left dislocated obliques and possessors also follow a pronominalisation strategy. I have nothing to say about this claim because I cannot understand her line of reasoning. For me, it is sufficient that S's and A's follow one pattern and O's follow another to say that the pattern of left dislocation is accusative.

Mosel (1987:470ff.) also argues that clauses with left dislocation have a structure similar to that of subject followed by a predicate. I would not want to call the dislocated nominal in these clauses a subject because virtually any clause level nominal can appear in that position and because the construction exhibits marked word order in that it is used to place contrastive focus on the dislocated nominal. Duranti (1981: section 8.2.3) expresses a similar opinion.

## 10. EQUI-NP DELETION

According to Keenan (1976:316), "the most likely NPs to undergo Equi-NP deletion include basic subjects". Thus the fact that a nominal undergoes equi-NP deletion (henceforth 'equi') suggests that it might be a basic subject. Chung (1978:106) describes equi as a rule which "deletes an NP of an embedded clause (the target) under coreference with some NP of the next higher clause (the controller)". In Samoan, the higher clause contains a verb of motion, volition, effort or persuasion (Chung 1978:125). Mosel (1987:468) claims that equi is "solely semantically determined. If the controlling NP functions as S in the matrix clause, it optionally triggers the deletion of the coreferential agentive participant in the embedded clause".

The type of equi that Mosel is referring to (S-controlled equi) needs to be broken down further into two types which differ depending on the nature of the matrix verb.<sup>12</sup> This verb is either one that takes an experiencer S, such as *mana'o* 'want', or a verb like *alu* 'go' that takes an agentive S. Let us first look at equi with verbs like *mana'o*, which I will refer to as '*mana'o* equi'.

With respect to this type of equi, Mosel's description is true if we only consider canonical transitive and intransitive verbs (and if we understand "agentive participant" broadly enough to include experiencers such as the primary participants of verbs of perception and cognition). The understood agentive S and A of the embedded clauses in (23a) and (23b) are affected by this rule, but the patientive O in (23c) is not.

- (23) a. E mana'o le tama e alu 'i Sāmoa.  
 IMP want the boy INF go DIR Samoa  
 The boy wants to go to Samoa.

<sup>12</sup>There are also other types of equi, for example, patient-controlled equi, as in "The girl was sent to wash the dishes." Mosel (1985:131-132) has shown that these other types of equi can have an S, an A, or an O as their target. Thus they argue for neither accusativity nor ergativity in Samoan.



- b. *E mana'o le tama e opo le teine.*  
 IMP want the boy INF hug the girl  
 The boy wants to hug the girl.
- c. *E mana'o le teine e opo \*(o ia) e le tama.*  
 IMP want the girl INF hug ABS she ERG the boy  
 The girl wants the boy to hug her.

However, if we consider the facts of *lavea* clauses, things are a little different. Sentences (24a) and (24b) show that the patientive S (but not the agent in the locative) of such a clause undergoes *mana'o* equi. Thus it must not be agents per se that are eligible for this rule.<sup>13</sup>

- (24) a. *E lē mana'o le teine e lavea i le tama.*  
 IMP NEG want the girl INF hurt LOC the boy  
 The girl doesn't want to be hurt by the boy.
- b. *\*E lē mana'o le tama e lavea le teine.*  
 IMP NEG want the boy INF hurt the girl  
 (The boy doesn't want to hurt the girl.)

The constraint on *mana'o* equi, I would claim, is the same as that on clitic placement: the target must be a non-zero S or A. Thus the S of *aulelei* 'be handsome' is not eligible for *mana'o* equi since it is a zero.

- (25) *E mana'o le tama e aulelei \*(o ia).*  
 IMP want the boy INF handsome ABS he  
 (The boy wants to be handsome.)

The argument based on *lavea* clauses would be weakened, of course, if the nominal in the locative in a *lavea* clause could be shown to be something other than an agent. Mosel (1985:14) claims that the participant in the locative in *lavea* clauses is not an agent performing an action but an "initiator" whose past action has brought about a state. *Lavea* verbs have traditionally been classified as 'stative verbs', and one would take that definition to mean that they designate states, but I am not convinced that they only (or even primarily) occur in clauses that indicate states.<sup>14</sup> I would claim that they also occur in clauses that designate processes, and in support of this claim I offer the following. Assuming, as illustrated in (26a) and (26b), that states readily co-occur with durative temporal expressions and processes with punctual ones, clauses like (26c) and (26d) indicate that *lavea* verbs can indicate processes.

- (26) a. *Na ma'i le tama i aso e tolu.*  
 PAST sick the boy LOC day IMP three  
 The boy was sick for three days.
- b. *Na fufulu e le tama le ta'avale*  
 PAST wash ERG the boy the car  
 The boy washed the car

<sup>13</sup>This argument is due to Chung (1976, 1978). Chung (1978: 131) marks a clause similar to (24a) as ungrammatical, which led her to say that the targets of equi needed to be both subjects and agents. My consultants do not find clauses like (24b) ungrammatical. Mosel (1987:468) was working with Chung's evaluation of that clause type.

<sup>14</sup>'Stative verb' is somewhat of a misnomer. Perhaps these verbs were first called 'stative' because they do not take an actor as their primary participant.

*i le ono i le afiafi.*  
 LOC the six LOC the evening  
 at six in the evening.

- c. \**Na lavea le tama i aso e tolu.*  
 PAST hurt the boy LOC day IMP three  
 ?The boy was hurt for three days.
- d. *Na lavea le tama i le ono i le afiafi.*  
 PAST hurt the boy LOC the six LOC the evening  
 The boy was hurt at six in the evening.

Even if it turns out that *lavea* verbs do designate only states, it doesn't seem impossible that agents might occur in a stative description. Consider, for example, the English sentence 'This book is written by John Steinbeck', which, I would claim, is stative but contains an agent. Also, if we could say that the boy in (2b) is responsible for the girl being hurt but he didn't *intend* to hurt the girl, the boy would be less than a canonical agent. However, consultants state that it is just as likely that the boy in (2b) hurt the girl intentionally.<sup>15</sup>

Let us now turn to *alu equi*, that is, *equi* with verbs like *alu* 'go' which take an agentive S. As illustrated in (27), this version of *equi* is limited to targets that are A's or agentive S's. O's cannot undergo *alu equi*.

- (27) a. *Na alu le tama e fufulu le ta'avale.*  
 PAST go the boy INF wash the car  
 The boy went to wash the car.
- b. *Na alu le tama e ta'alo i le paka.*  
 PAST go the boy INF play LOC the park  
 The boy went to play in the park.
- c. \**Na alu le teine e opo e le tama.*  
 PAST go the girl INF hug ERG the boy  
 (The girl went to be hugged by the boy.)

As shown in (28a) and (28b), the S and locative agent of *lavea* clauses are not eligible for *alu equi*. The S is ruled out because it is not an agent, and the locative agent because it is not an S or A.

- (28) a. \**Na alu le teine e lavea i le tama.*  
 PAST go the girl INF hurt LOC the boy  
 (?The girl went to be hurt by the boy.)
- b. \**Na alu le tama e lavea le teine.*  
 PAST go the boy INF hurt the girl  
 (?The boy went to hurt the girl.)

Mosel (1987:468) claims that *equi* is "solely semantically determined" in that it is limited to agents. Although it is true that the target of *alu equi* must be an agent, it is not true that all agents undergo this rule. Sentence (28b) shows that the agent of a *lavea* clause does not undergo *alu equi*. Since this rule is restricted to agents that are S's or A's, it provides further evidence for accusativity in Samoan.

<sup>15</sup>There is a connotation that the offence involved in sentence (2b) is a sexual one.

## 11. RAISING

Keenan (1976:320) states that "basic subjects are always among the NPs in a language that can undergo raising". Thus nominals that undergo raising are likely to be subjects. Chung (1976, 1978) describes two versions of raising in Samoan: one involves predicates like *māsani* 'be usual, be used to' and the other, predicates such as *mafai* 'can'. *Māsani* raising affects clause-level NPs of all syntactic types (Chung 1978:202); therefore, it does not argue for accusativity or ergativity, and so I will not discuss it here. *Mafai* raising, however, is limited to certain nominal types. This rule relates sentences like (29a) and (29b).

- (29) a. *E mafai ona tautala le pepe.*  
 IMP can COMP talk the baby  
 The baby can talk.
- b. *E mafai e le pepe ona tautala.*  
 IMP can ERG the baby COMP talk  
 The baby can talk.

Sentences (29a) and (29b) illustrate the fact that *mafai* raising affects agentive S's. Sentences (30a) and (30b) show that A's (but not O's) can be targets of *mafai* raising.

- (30) a. *E mafai e le tama ona fufulu le ta'avale.*  
 IMP can ERG the boy COMP wash the car  
 The boy can wash the car.
- b. *\*E mafai (e) le teine ona opo e le tama.*  
 IMP can ERG the girl COMP hug ERG the boy  
 (The girl can be hugged by the boy.)

*Mafai* raising, however, does not affect S's that are patients nor agents that are not S's or A's. Thus, the S of a *lavea* clause fails to undergo raising because it is not agentive, and the agent of such a clause also fails to undergo this rule because it is neither an S nor an A.

- (31) a. *\*E mafai (e) le teine ona lavea i le tama.*  
 IMP can ERG the girl COMP hurt LOC the boy  
 (The girl can be hurt by the boy.)
- b. *\*E mafai (e) le tama ona lavea le teine.*  
 IMP can ERG the boy COMP hurt the girl  
 (The boy can hurt the girl.)

Since the agent of a *lavea* verb does not undergo *mafai* raising, it is not exactly agents that are affected by this rule. Instead, it is agents that are S's or A's. The fact that *mafai* raising needs to refer to S's and A's at all is still another indication of accusativity in Samoan.

## 12. IMPERATIVES

Another subject characteristic mentioned by Keenan (1976:321) is that of being the addressee of an imperative. This phenomenon also reveals accusativity in Samoan in that the addressees of imperatives are agents that are either S's or A's:

- (32) a. *Alu 'i le fale.*  
 go DIR the house  
 Go to the house (i.e. go home).
- b. *Fufulu le ta'avale.*  
 wash the car  
 Wash the car.

Notice that it is not agency per se that allows a nominal to serve as the addressee of an imperative. This is evidenced by the fact that the agent of a *lavea* clause cannot be the addressee of an imperative because it is not an S or an A.

- (33) \**Lavea le mogamoga.*  
 hurt the cockroach  
 (?Hurt the cockroach.)

### 13. THE CODING OF CAUSERS

Keenan (1976:321) notes that basic subjects normally have the same coding properties (case marking, etc.) as causers in basic causative clauses. In Samoan, the prefix *fa'a-* derives causative verbs. Some causative verbs are intransitive (i.e. they involve self-induced causation), a few are middle, but the majority are transitive. Sentences (34a)-(34c) illustrate each of these types. The point of interest here is that the causer in the intransitive and middle clauses is an S, while the causer in the transitive clause is an A. (In the case of intransitive causatives, the causer is identical to the causee.) The coding of causers provides further evidence of accusativity in Samoan in that it conflates S's and A's.

- (34) a. *Na fa'a-māvae le uō i le tāulaga.*  
 PAST CAUS-part the friend LOC the harbor  
 The (two) friends parted at the harbor. (Milner 1966:142)
- b. *'Olo'o fa'a-logo le tama 'i le faiā 'oga.*  
 PROG CAUS-inform the boy DIR the teacher  
 The boy is listening to the teacher.
- c. *Na fa'a-inu e le tama le solofanua.*  
 PAST CAUS-drink ERG the boy the horse  
 The boy watered the horse. (lit. The boy caused the horse to drink.)

### 14. SUBJECT AS A CONFLATION OF S AND A

The above observations can be summarised as follows. Case marking, traditional verb agreement, and relativisation follow an ergative pattern: these phenomena group together S's and O's at the exclusion of A's. The phenomena that show an accusative pattern are dispensability, word order, clitic placement, left dislocation, *mana'o* and *alu* equi, *mafai* raising, imperatives, and the coding of causers. These phenomena affect A's and S's (and not O's). Deletion across conjunctions follows neither an ergative nor an accusative pattern. Speaking in numerical terms, on the basis of what we

have seen so far, Samoan is clearly more accusative than ergative in view of the fact that there are three times as many accusative as ergative phenomena (three ergative and nine accusative).<sup>16</sup>

If we follow Dixon (1979:108) in defining 'subject' as the conflation of S and A, then our evidence for accusativity is also evidence for the category subject. The phenomena that group together S's and A's in Samoan indeed argue for this category. This is not to say, however, that there is not also evidence for the category absolutive in Samoan. The phenomena that unite S's and O's also argue for this second category.

## 15. EVIDENCE FOR DEEP AND SURFACE SUBJECTS

According to Dixon (1979:109), the definition of subject as the conflation of S and A should be considered that of deep subjects. The question then arises as to whether there is any evidence in Samoan for surface subjects that are not also deep subjects or deep subjects that are not also surface.<sup>17</sup> I believe there is such evidence. Since an A is agentlike, I will assume that it is a deep subject no matter where it appears in a clause. However, I will argue that an A is a *surface* subject in a VAO but not in a VOA clause. This claim is supported by the fact that an A has certain subject properties when it occurs in VAO but not VOA clauses. The relevant phenomena are leftmost position, quantifier float, reflexivisation and modern verb agreement.

## 16. LEFTMOST POSITION

I argued above that the basic word order of transitive clauses is VAO and noted that one of Keenan's subject properties, that of being leftmost among nominals, suggests that the A of a VAO clause is a subject. As illustrated in (35), an A has the subject property of being leftmost in a VAO but not a VOA clause.

- (35) a. *Na fufulu e le tama le ta'avale.*  
 PAST wash ERG the boy the car  
 The boy washed the car.
- b. *Na fufulu le ta'avale e le tama.*  
 PAST wash the car ERG the boy  
 The car was washed by the boy.

## 17. QUANTIFIER FLOAT

The ability to launch a floating quantifier is also among Keenan's subject characteristics (1976:320). In Samoan, a rule which Chung (1976:194-198) refers to as "quantifier float" moves the

<sup>16</sup>Keenan (1976) proposes about 30 subject properties. Ideally we would go through them one by one for Samoan, but such a project is beyond the scope of this paper. My impression is that the statistical balance would not change greatly even if we were to consider all 30 properties.

<sup>17</sup>*The car* and *the boy* in the English passive sentence *The car was washed by the boy* respectively exemplify a surface subject that is not a deep subject and a deep subject that is not a surface subject. The terms 'deep subject' and 'surface subject' are, of course, from transformational grammar. Other theories have made similar distinctions in types of subjects, for example, initial and final subject in relational grammar and logical and grammatical subject in traditional grammar.



quantifier 'uma from an NP and places it right after the verb. Quantifier float relates clauses like (36a) and (36b).

- (36) a. 'Ua ō tamaiti 'uma 'i Sāmoa.  
 PERF go(PL) children all DIR Samoa  
 All the children have gone to Samoa.
- b. 'Ua ō 'uma tamaiti 'i Sāmoa.  
 PERF go(PL) all children DIR Samoa  
 The children have all gone to Samoa.

In my own work on this phenomenon, I have found that S's and O's can be bound by postverbal 'uma no matter where they occur in a clause, while A's must be leftmost among the postverbal nominals in order to be bound.<sup>18</sup> Sentences (37a) and (37b) show that S's in clitic position can be bound by 'uma, but A's in that position cannot. Sentences (38) and (39) illustrate that O's can be bound in any position, but A's must be leftmost.

- (37) a. 'Ua lātou ō 'uma 'i Sāmoa.  
 PERF they go(PL) all DIR Samoa  
 They all have gone to Samoa.
- b. \*Lātou te uli 'uma la'u ta'avale.  
 they IMP drive all my car  
 (They all drive my car.)
- (38) a. Na 'ave 'uma e le tama tusi.<sup>19</sup>  
 PAST take all ERG the boy book  
 The boy took all the books.
- b. Na 'ave 'uma tusi e le tama.  
 PAST take all book ERG the boy  
 The books were all taken by the boy.
- (39) a. E uli 'uma e a'u uō la'u ta'avale.  
 IMP drive all ERG my friend my car  
 My friends all drive my car.
- b. \*E uli 'uma la'u ta'avale e a'u uō.<sup>20</sup>  
 IMP drive all my car ERG my friend  
 (My car is driven by all my friends.)

Given that S's and O's behave similarly in that they both can be bound by 'uma when they are in any position (while A's need to be in immediate postverbal position), quantifier float without the word order restriction (henceforth unrestricted quantifier float) can be said to be organised in an

<sup>18</sup>Certain 'obliques' can also trigger quantifier float. See Cook (1988:47, 84) for details.

<sup>19</sup>As can be anticipated, if both the ergative and the absolutive are plural in a clause with VAO word order, ambiguity results:

- (i) Na 'ave 'uma e tamaiti tusi.  
 PAST take all ERG children books  
 The children all took books. or  
 The children took all the books.

<sup>20</sup>Barbara Voigt (pers. comm.) reports that her Western Samoan informants accept clauses like (39b). There may be dialectal differences here, given that most of my informants are from American Samoa.

ergative manner. However, what is of interest here is the fact that A's have the subject property of 'launching a floating quantifier' when they are leftmost, that is when they are in VAO but not VOA clauses.

## 18. REFLEXIVISATION

Keenan (1976:315) states that "in some languages control of reflexives within clauses is largely restricted to basic subjects". According to Chapin (1970), either the A or the O of a Samoan transitive clause can control reflexivisation as long as the antecedent precedes the pronoun.<sup>21</sup>

- (40) a. *E viv'i e le tama ia lava.*  
 IMP praise ERG the boy he self  
 The boy praises himself.
- b. *E viv'i le tama e ia lava.*  
 IMP praise the boy ERG he self  
 The boy is praised by himself.
- c. *\*E viv'i e ia lava le tama.*  
 IMP praise ERG he self the boy  
 (\*He<sub>i</sub> praises the boy<sub>i</sub>.)
- d. *\*E viv'i 'o ia lava e le tama.*  
 IMP praise ABS he self ERG the boy  
 (\*He<sub>i</sub> is praised by the boy<sub>i</sub>.)

At first glance, it would seem that the only thing that matters here is that the antecedent precede the pronoun. However, if we look also at middle clauses, we see that the situation is more complicated than that. As illustrated in (41), S's (but not obliques) also control reflexivisation. Again the antecedent must precede the pronoun.

- (41) a. *E alofa le tama 'iate ia lava.*  
 IMP love the boy DIR he self  
 The boy loves himself.
- b. *\*E alofa 'i le tama 'o ia lava.*  
 IMP love DIR the boy ABS he self  
 (\*He<sub>i</sub> loves the boy<sub>i</sub>.)
- c. *\*E alofa 'o ia 'i le tama.*  
 IMP love ABS he DIR the self  
 (\*He<sub>i</sub> loves the boy<sub>i</sub>.)
- d. *\*E alofa 'iate ia lava le tama.*  
 IMP love DIR he self the boy  
 The boy loves himself.

<sup>21</sup>Chapin's (1970) original sentences contained the verb *sogi* 'cut/chop' rather than *viv'i* 'praise'. Some consultants reject *sogi* in such sentences, but will accept *viv'i*. When one talks about bodily wounds in Samoan, one usually mentions the body part that is affected, as in 'the boy cut his arm' rather than 'the boy cut himself'. Praising, however, can refer to the individual as a whole.

What is of immediate interest with respect to this rule is the fact that the A of a VAO clause (but not that of a VOA clause) has the subject property of controlling reflexivisation.

### 19. MODERN VERB AGREEMENT

As we saw above, Mosel's characterisation of verb agreement, which I labelled 'traditional verb agreement', is that if a verb agrees in number with some nominal, it agrees with the absolutive. My observation is that the situation is slightly more complicated. Apparently I am not the only one holding this view. Churchward (1951:78) points out that speakers use singular forms for plural and vice versa, and Marsack (1962:36) observes "a growing tendency to use the same form for both singular and plural". Our main concern here is with transitive verbs. Although, as Mosel (1987:457) points out, there are many examples of absolutives agreeing with transitive verbs in published texts, the texts that she mentions were transcribed around the turn of the century (Krämer 1902-03; Stuebel 1896).

In Hovdhaugen (1987), a more current collection of (seven) oral texts, verb agreement with transitive verbs is rather rare. The transitive verbs given in (42) are the only ones that appear with plural forms in the glossary, which supposedly includes all the words that occur in the 19 or so pages of the texts. Example (42) also includes the plural form *sāisai*, which appears in one text (p.46), but not in the glossary.

(42)	Singular	Plural	
	<i>fa'a'a'e</i>	<i>fa'afe'a'ei</i>	cause to climb
	<i>fasi</i>	<i>fafasi</i>	beat, kill (an animal)
	<i>lagona</i>	<i>lāgolagona</i>	feel, perceive, sense
	<i>saisai</i>	<i>sāisai</i>	bind
	<i>su'e</i>	<i>sū'e</i>	look for, search
	<i>tāumafa</i>	<i>tāumamafa</i>	eat (polite word)

Not only are the plural forms of transitive verbs relatively rare in Hovdhaugen, but two of those plural forms given in (42) are not even used as transitive verbs. Instead they are used as middle or intransitive verbs. This is true for *sū'e* (pp.54, 198), and *tāumafa* (p.205).

Milner's (1966) dictionary gives plural forms for several transitive verbs but not many example sentences in which it is clear that the verb agrees with the absolutive but not the ergative. In fact, he gives at least one sentence (an infinitive in an imperative) in which the verb agrees with the ergative (the addressee) and not the absolutive. (The *lau* referred to in (43) is part of a communal fishing device.)

(43)	<i>Ō</i>	<i>mai</i>	<i>e</i>	<i>tata'i</i>	<i>le</i>	<i>lau.</i>
	come(PL)	hence	INF	draw(PL)	the	<i>lau</i>
	Come and draw the <i>lau</i> . (Milner 1966:230)					

When presented with the transitive verb forms given in Milner (1966), my informants responded that they do not use many of them. Furthermore, when asked to choose between the singular and plural forms with which they were familiar in sentences with verbs left out, they would most frequently make the verb agree with the absolutive, but they would also, upon occasion, produce the verb agreement pattern that appears in (44). This pattern parallels that of quantifier float. The verb agrees with the O no matter where it occurs in the clause, and it agrees with the A of a VAO clause. (I have

omitted here the cases of only singular or only plural nouns, in which cases the verb chosen was respectively singular or plural.) Least frequently, my informants would make the verb agree only with the nominal following the verb, whether this was an O or an A. This latter pattern is similar to that of reflexivisation in that the verb agrees with the A or the O, whichever occurs leftmost. I will refer to these two non-traditional patterns of verb agreement as 'modern verb agreement'.

- (44) a. *Na tutuli e ali'i le 'avefe'au.*  
 PAST send(PL) ERG chief the messenger  
 The chiefs sent the messenger.
- b. *Na tuli (\*tutuli) le 'avefe'au e ali'i.*  
 PAST send (send(PL)) the messenger ERG chief  
 The messenger was sent by the chiefs.
- c. *Na tutuli e le ali'i 'avefe'au.*  
 PAST send(PL) ERG the chief messenger  
 The chief sent the messengers.
- d. *Na tutuli 'avefe'au e le ali'i.*  
 PAST send(PL) messenger ERG the chief  
 The messengers were sent by the chief.

I take this variation concerning verb agreement to be an indication that the use of plural transitive forms is dying out. However, of special interest is the fact that in the cases in which the verb did not agree solely with the O, it agreed with the A in leftmost position. Thus in VAO clauses, once again, the A manifests a subject property which it does not have in VOA clauses.

## 20. THE GRAMMATICAL RELATIONS OF A AND O

In sum, the A of a VAO clause manifests four more subject properties than the A of a VOA clause. It is leftmost and it triggers quantifier float, reflexivisation and modern verb agreement. Assuming that subjects that are both deep and surface subjects manifest more subject properties than those that are not, and given that an A manifests the semantics of a deep subject (i.e. it is an agent), I propose that the A of a VAO clause is both deep and surface subject and that that of a VOA clause is a deep subject but not a surface subject.

With respect to the O in each clause type, I suggest that since the O of a VAO clause and the S of an intransitive clause are both absolutes and they both share the properties of case marking, traditional verb agreement, relativisation and unrestricted quantifier float (exactly those properties that the A does not have), there is no need to analyse the O of a VAO clause as anything other than an absolute.<sup>22</sup> The O of a VOA clause, in contrast, has the added features of being leftmost, controlling reflexivisation, and controlling modern verb agreement when the verb agrees only with the leftmost nominal. Although this evidence is not overwhelming, it suggests that the O of a VOA clause is a surface (but not deep) subject, and so I will tentatively analyse it as such.

<sup>22</sup>Chung (1978:183-188) posits a rule of object incorporation, which, she claims, argues for the category of direct object in Samoan. Since this rule has lexical exceptions and applies to some obliques as well as absolutes, I would analyse it as a lexical (rather than syntactic) rule that derives intransitive verbs from transitive.

## 21. THE STRUCTURE OF TRANSITIVE CLAUSES

Given the above analyses of A's and O's, a VAO clause is similar to an active transitive clause in more familiar languages (like English) except that it contains an absolutive rather than a direct object. As for VOA clauses, since the A is a deep (but not surface) subject and the O is a surface (but not deep) subject, I will analyse such clauses as passive.<sup>23</sup>

Linguists accustomed to seeing special verbal morphology associated with passive will probably find it strange that I analyse VOA clauses as passive.<sup>24</sup> I do not find it logically impossible that a language might show an active/passive contrast with only word order inversion (i.e. inversion of the A and O). Perlmutter and Postal (1983) and Siewierska (1984) have shown that among the devices of word order inversion, special morphology, and special case marking, there is no single device that languages use to indicate passive. In fact, Siewierska (1984:24) states that word order inversion is "characteristic of the passive in many languages". In order to arrive at a cross-linguistic generalisation concerning the passive construction, Siewierska (p.75) finds it necessary to characterise this construction in terms of the corresponding active. In particular, she states (p.256) that "the event or action expressed in the passive is brought about by some person or thing which is not the passive subject, but the subject of the corresponding active". If we understand 'subject' here as 'surface subject', the present analysis of VAO and VOA clauses as respectively active and passive fits this characterisation.

One may also ask how the case marking of VAO and VOA clauses is to be accommodated if these clauses are respectively active and passive. Case marking, I would claim, indicates deep relations for both clause types. Whether the clause is active or passive, the (deep) transitive subject is in the ergative and the (deep) absolutive is the absolutive case. Likewise the phenomena that refer to subject as the conflation of S and A refer to the A as a deep subject, and the phenomena that require that a nominal be leftmost (i.e. in a VAO clause) refer to surface subject. As for quantifier float and the version of modern verb agreement that parallels quantifier float, these phenomena affect both surface subjects and absolutives.

## 22. SUMMARY AND CONCLUSION

In this paper I have shown that there is evidence for subject defined as the conflation of S and A. The phenomena of dispensability, word order, clitic placement, left dislocation, *mana'o* and *alu equi*, *mafai* raising, imperatives, and the coding of causers argue for this category. There is also evidence from case marking, traditional verb agreement, relativisation and unrestricted quantifier float for the category absolutive.

Following Dixon's (1979:109) proposal that the conflation of S and A be considered the definition of 'deep subject', I have assumed that the above-mentioned phenomena that conflate S's and A's affect A's because they are deep subjects. I have also assumed (on the basis of its semantic role of agent) that an A is a deep subject no matter where it occurs in a clause. Given these assumptions, I have shown that there is evidence for surface subjects that are not deep and deep subjects that are not surface. In particular, I have argued that the phenomena of quantifier float, reflexivisation, modern

<sup>23</sup>Even if it turns out that the O of a VOA clause is not a surface subject, the clause type can still be analysed as impersonal passive, that is, a passive in which neither the A nor the O is a surface subject.

<sup>24</sup>See Cook (1988:138-139) for reasons why the *Cia* suffix which marks passive in some other Polynesian languages should not be considered a passive suffix in Samoan. Neither Chung (1976,1978) nor Mosel (1987) considers the possibility that VAO and VOA differ structurally.



verb agreement, and leftmost position show that an A in a VAO clause is a surface subject, while the A in a VOA clause is not.

Finally, I have proposed that VAO clauses are similar to active transitive clauses in more familiar languages except that the O that they contain is an absolutive rather than a direct object. I have also analysed VOA clauses as passive although the evidence that the O of such a clause is a subject is not overwhelming. Hopefully the ongoing search for subject in Samoan will turn up more clues concerning the role of the O in VOA clauses.

## BIBLIOGRAPHY

- BIGGS, Bruce**, 1974, Some problems of Polynesian grammar. *Journal of the Polynesian Society* 83:401-426.
- CHAPIN, Paul G.**, 1970, Samoan pronominalization. *Language* 46:366-378.
- CHUNG, Sandra**, 1976, Case marking and grammatical relations in Polynesian. PhD dissertation, Harvard University.
- 1978, *Case marking and grammatical relations in Polynesian*. Austin: University of Texas Press.
- CHURCHWARD, Spencer**, 1951, *A Samoan grammar*. 2nd edn. Melbourne: Methodist Church of Australasia.
- COOK, Kenneth W.**, 1978, The mysterious Samoan transitive suffix. Proceedings of the Annual Meeting of the Berkeley Linguistics Society 4:53-66.
- 1988, A cognitive analysis of grammatical relations, case and transitivity in Samoan. PhD dissertation, University of California, San Diego.
- DIXON, Robert M.W.**, 1979, Ergativity. *Language* 55:59-138.
- DURANTI, Alessandro**, 1981, *The Samoan fono: a sociolinguistic study*. PL, B-80.
- HOVDHAUGEN, Even**, 1987, *From the land of Nāfanua: Samoan oral texts in transcription with translation, notes and vocabulary*. Oslo: Norwegian University Press.
- JOHNSON, Alan and Lillian HARMON**, n.d., *Let's speak Samoan*. Auckland: Church of Jesus Christ of Latter-Day Saints.
- KEENAN, Edward L.**, 1976, Towards a universal definition of 'subject'. In Charles N. Li, ed. *Subject and topic*, 304-333. New York: Academic Press.
- KRÄMER, Augustin F.**, 1902-03, *Die Samoa Inseln*. 2 vols. Stuttgart: E. Schweizerbartsche Verlagbuchhandlung.
- LANGACKER, Ronald W.**, 1987, *Transitivity, case, and grammatical relations: a cognitive grammar prospectus*. Linguistic Agency, University of Duisburg.
- MARSACK, C. C.**, 1962, *Teach yourself Samoan*. London: English Universities Press.
- MILNER, George, B.**, 1966, *Samoan dictionary*. London: Oxford University Press.
- 1973, It is aspect (not voice) which is marked in Samoan. *Oceanic Linguistics* 12:621-639.
- MOSEL, Ulrike**, 1985, Ergativity in Samoan. AKUP (Arbeiten des Kölner Universalien-Projekts) 61. Köln: Institut für Sprachwissenschaft Universität zu Köln.
- 1987, Subject in Samoan. In Donald C. Laycock and Werner Winter, eds *A world of language: papers presented to Professor S.A. Wurm on his 65th Birthday*, 455-479. PL, C-100.
- 1989, On the classification of verbs and verbal clauses in Samoan. In Ray Harlow and Robin Hooper, eds *VICAL 1: Oceanic languages: papers from the Fifth International Conference on Austronesian Linguistics*, 377-398. Auckland: Linguistic Society of New Zealand.
- MOYLE, Richard**, 1981, *Fagogo: fables from Samoa in Samoan and English*. Auckland: University Press.
- OCHS, Elinor**, 1982, Ergativity and word order in Samoan child language. *Language* 58:646-671.
- KEENAN, Edward L.**, 1976, Towards a universal definition of 'subject'. In Charles N. Li, ed. *Subject and topic*, 304-333. New York: Academic Press.

- PERLMUTTER, David and Paul POSTAL, 1983, Toward a universal characterization of passivization. In David Perlmutter, ed. *Studies in Relational Grammar 1*: 3-29. Chicago: Chicago University Press (original version in Proceedings of the Annual Meeting of the Berkeley Linguistics Society 3:395-417).
- SIEWIERSKA, Anna, 1984, *The passive: a comparative linguistic analysis*. London: Croom Helm.
- STUEBEL, C., 1896, *Samoanische Texte*. Berlin: Veröffentlichungen aus dem Königlichen Museum für Völkerkunde.
- TUI TELE, M.T., M.Z. SÂPOLU and J. KNEUBUHL, 1978, *Lâ Tâtou Gagana: Tusi Muamua* [Our Language: Book I]. Pago Pago: Bilingual/Bicultural Education Project of American Samoa.
- VOIGT, Barbara, 1986, 2-To-3 retreat in Samoan. Working Papers in Linguistics, University of Hawaii, 17:77-116.

## SHIFTING ATTITUDES IN HAWAII 1920s TO 1990

SAMUEL H. ELBERT

The American missionaries to Hawai'i in the 1820s embraced the local language. They realised that the way to the heart of another culture was through the language. They had studied Latin and Greek and some knew Hebrew, so they had some idea of the cost in time and concentration in mastery of another language, especially one as different from English and the three classical languages as is Hawaiian. At first they wisely considered an alphabet. Fortunately they were aided by William Ellis, a British missionary who had served in Tahiti with the London Missionary Society, and with his guidance the Americans devised an alphabet consisting of five vowels and seven consonants that was so easy to learn that it survived intact for about 150 years.

Nevertheless, once the Hawaiians were literate and learning English, a campaign against the use of Hawaiian was waged by government, public and private schools (especially the Kamehameha Schools), and by Hawaiian parents who scolded if their children answered their Hawaiian in Hawaiian. "Speak English!" they would say.

A little girl who turned out eventually to be a famous Hawaiian scholar was punished for whispering a few Hawaiian words to classmates from the country who couldn't understand their teacher's English. This was Mary Kawena Pukui and the place was a school in Mānoa Valley run by Kawaiaha'o Church, now considered a refuge for Hawaiiana. For translating a few words she was made to stand in the corner of the classroom. Much worse was her punishment for explaining what a napkin was used for. For seven days she had to take her meals alone in the centre of the room, and her 'food' consisted of bread and water and no *poi*. She was also scolded for immorality; she had danced a few *hula* steps. Nevertheless when she grew up and worked at the Bishop Museum, she befriended the Reverend Henry H. Parker of Kawaiaha'o Church. She and her mother helped him with his revision of *A dictionary of the Hawaiian language* by Lorrin Andrews (1865).

Kawena always loved the skimpy Hawaiian alphabet, and decades later wrote the words and music for what she called an 'alphabet song' for her first grandchild, La'akea (reprinted in Elbert 1970):

*E nā hoa kamali'i*  
*E a'o mai kākou*  
*I pa'ana'au ka pī'āpā*  
*Ā, 'ē, 'ī, 'ō, 'ū*

O fellow children  
Let us learn together  
Till we've memorized the alphabet  
A, e, i, o, u

Hō, kō, lā, mū, nū  
 'O pī me wō nā panina  
 O ka pī'āpā

H, k, l, m, n  
 P and w are the last  
 Of the pī'a'pā

Note that in the Hawaiian original, an apostrophe and a line over some vowels occurred as additions to the five vowels and seven consonants mentioned in the song. These same additional symbols had been used by Kawena for many years in her scholarly articles. They are discussed later.

The Hawaiian language had been taught for many years at the Mānoa campus by highly respected ministers, fluent speakers of the language who gave beautiful Hawaiian prayers at commencement exercises and other ceremonies, but who knew nothing about the nature of language (who does know?) or about how to teach a language. Their grading systems were filled with compassion. Students who came frequently to classes received A, students who seldom came received B, and those who registered but never appeared received 'Incomplete'. This appealed to football players. One instructor boasted that he won more games for the University of Hawai'i than any other instructor. Members of the English Department were not amused and advised the better students to stay out of Hawaiian classes. Too much *aloha*, they thought.

In the 1960s, however, attitudes towards Hawaiian were changing. Those known as Chinese-Hawaiians began to call themselves Hawaiian-Chinese or just Hawaiian. Young people were embarrassed that they could not speak the mother tongue. At the University students learned that confusing pairs, trios and quatrains that were written in the same way, could actually be written to show that the words were not at all the same. Thus:

- (a) *mai* 'this way'; *ma'i* 'sick, genital'
- (b) *ai* 'to have sex'; *ai* 'to eat'; *ā'ī* 'neck'
- (c) *pau* 'finished'; *pa'u* 'drudgery'; *pā'ū* 'moist'; *pā'ū* 'sarong'

And the English professors decided that Hawaiian was a beautiful language with a rich literature, and a master's degree was earned by Derryl Cabacungan for writing poetry and an opera in Hawaiian with noteworthy English translations. The State legislature repealed the 1896 law against the use of Hawaiian in public schools and voted to make Hawaiian an official language of the State of Hawai'i.

Meanwhile University students and others were thinking that Dr Pukui (she had two doctorates) had been right about the two additions to the alphabet. Two young teachers of Hawaiian led the campaign. They were Larry Kimura and Pila Wilson, both fluent speakers of Hawaiian who knew only too well that many students of Hawaiian ancestry in their classes did not know how to pronounce the word for good (now written *maika'i*) and the name of the State song, 'Hawai'i Pono'ī'. The better students and the leaders insisted that the two new symbols be adopted by everyone who wrote Hawaiian words.

But what to call these names? The apostrophe, sometimes reversed and sometimes not, had long been known to linguists as glottal stop. It had also been known in Hawai'i as a *hamzah* and *'u'ina*. Glottal stop was too scholarly and *hamzah* was Arabic. The old Hawaiian name *'u'ina* meant 'to crackle', hardly appropriate. Kimura and Wilson liked the term used by Elbert in *Spoken Hawaiian*, *'okina* 'break'. The term *'okina* was accepted quickly, with gratitude by many, and was printed in newspapers and on maps, but many scientists complained that it was idiosyncratic and not scholarly, and implied that it was an expensive nuisance to print.

The other *malihini* addition had long been known to scholars as a macron, but this came from Greek. It too should have a Hawaiian name, Kimura and Wilson said, and they introduced the term

*kahakō* 'long mark'. The authors, however, did not admit that in final position the stress is the identifying principle, as here the length is much less noticeable than the stress. Ordinary unmarked vowels in final position are often whispered or omitted, but this never happens to vowels with *kahakō*.

The following table shows how the two additions to the alphabet were accepted. The works cited were published by the Bishop Museum and the University of Hawai'i Press.

		'okina	kahakō
1943	Judd, Pukui and Stokes	yes	yes
1951	Malo	no	no
1952	Titcomb and Pukui	yes	yes
1957	Buck	yes	no
	Pukui and Elbert	yes	yes
1958	Handy and Pukui	yes	no
1959	Ii, Pukui and Barrère	yes	no
1964	Kamakau, Ka po'e kahiko, Pukui and Barrère	yes	no
1971	Malo	no	no
1976	Kamakau, <i>The works ...</i> , Pukui and Barrère	yes	no

Only the first Pukui and Elbert work is listed. Note that Pukui, as usual, was way ahead of her time, and that the Bishop Museum was very slow to accept the inevitable. An outstanding exception was the City and County of Honolulu. Street names are now being written with the new symbols. Tourists do not know what they mean, or how they affect the pronunciation, but they do not complain. Hawai'i was too Americanised.

## FINALE

Our honoree has not been named in this essay, but his knowledge of the Pacific is so vast that he has even written about Hawaiian. I, Samuel Elbert, first met him when he passed through Honolulu in 1955 on his way to the South Pacific to study the Malayo-Polynesian language family everywhere, especially in Melanesia and Indonesia. Several years later Howard McKaughan told me that George Grace was interested in coming to Hawai'i. What did I think about it? "Wonderful!" I said enthusiastically and I say this again!

## ACKNOWLEDGEMENTS

I am grateful to William (Pila) H. Wilson and Michel Kaiser for critical reading of early drafts of this article.

## BIBLIOGRAPHY

BUCK, Peter H. See Hiroa.

ELBERT, Samuel H., 1970, *Spoken Hawaiian*, illustrated by Jean Charlot. Honolulu: The University of Hawaii Press.

HANDY, E. S. Craighill and Mary Kawena PUKUI, 1958, *The Polynesian family system in Ka-'u, Hawai'i*. Wellington: The Polynesian Society.



- HIROA, Te Rangi (Sir Peter H. BUCK)**, 1957, *Arts and crafts of Hawaii*. Honolulu: Bishop Museum Press, Special Publication 45.
- II, John Papa**, 1959, *Fragments of Hawaiian history*, as recorded by J.P. Ii; translated by M.K. Pukui, edited by D.B. Barrère. Honolulu: Bernice P. Bishop Museum Press.
- JUDD, Henry P., Mary Kawena PUKUI and John F. G. STOKES**, 1943, *Introduction to the Hawaiian language (an English-Hawaiian vocabulary); with a complementary Hawaiian-English vocabulary*. Honolulu: Tongg.
- KAMAKAU, Samuel Manaiakalani**, 1964, *Ka po'e kahiko, the people of old*; translated from the newspaper *Ke Au 'Okō'a* by Mary Kawena Pukui; arranged and edited by Dorothy B. Barrère. Honolulu: Bishop Museum Press, Bernice P. Bishop Special Publication 51.
- 1976, *The works of the people of old, Na hana aka po'e kahiko*; translated from the newspaper *Ke Au 'Okō'a* by Mary Kawena Pukui; arranged and edited by Dorothy B. Barrère. Honolulu: Bishop Museum Press, Bernice P. Bishop Museum Special Publication 61.
- MALO, David**, 1951, 1971, *Hawaiian antiquities (Moololo Hawaii)*; translated from the Hawaiian by Dr. Nathaniel B. Emerson 1898. Honolulu: Bishop Museum Press, Special Publication 2.
- PUKUI, Mary Kawena and Samuel H. ELBERT**, 1957, *Hawaiian dictionary: Hawaiian-English, English-Hawaiian*. Honolulu; University of Hawaii Press.
- TITCOMB, Margaret, with the collaboration of Mary Kawena PUKUI**, 1952, *Native use of fish in Hawaii*. New Plymouth: Avery Press. (1st edn 1940, Polynesian Society Memoir 29.)

## DETERMINER FORMS IN HOTE

ELAINE M. GOOD

### 1. INTRODUCTION

Hote is an Austronesian language spoken in the Morobe Province of Papua New Guinea by about 3,200 people (Muzzey 1988). Traditionally, Austronesian linguists have considered Hote to be an isolate, perhaps a separate family within a 'Siassi family' (Hooley 1971). More recently, Ross (1988:133) lists Hote as part of a South Huon Gulf chain in the Huon Gulf family and defends the chaining of Kaiwa, Hote and Buang that was earlier implied by Bradshaw (1978).

This paper is part of a more thorough grammatical study which is now in progress. The initial grammatical and phonological analysis of Hote was done by Marguerite Muzzey (1979, 1988). Since this paper is part of initial research on this language, it first addresses the question: what are the determiners in Hote? Muzzey considers that six forms in the language may be determiners because of their occurrence at the end of noun phrases (pers. comm.). However, other data must be taken into account to decide what forms are actually determiners. As a working definition, I will define determiners as forms that occur with noun phrases and specify, or limit, their semantic range in some identifiable manner.

The material in this research comes mainly from about 125 pages of transcribed texts.<sup>1</sup> These texts were tape-recorded from various speakers of the Hote dialect, apparently the most central of three dialects (Misim, Yamap and Hote; cf. Ross 1988:134, note 55), then transcribed and translated by Muzzey.

I begin by giving a short overview in section 2 of Hote grammar to facilitate the reader's understanding of the following discussion.

---

<sup>1</sup>I have also spent some time in the village of Hote, working on the material and learning the language. However, this time is still too short to have provided a good database in itself.

## 2. OVERVIEW OF HOTE GRAMMAR

A declarative independent clause has the following order in Hote:

Time, *ming*<sup>2</sup>. Subj.NP, V<sub>1</sub> (V<sub>2</sub>...V<sub>n</sub>), Adv, DO NP, Particle, Location, Time, *ami*

An I(ndirect) O(bject) Noun Phrase may occur either before or after the D(irect) O(bject) NP. Location and Time may be either prepositional phrases or noun phrases. Particles are analogous to those noted for such English phrases as 'sweep it **up**'. Serial verbs, indicated above by repetition of V with ascending numbers, are the rule rather than the exception in Hote. Forms in the C(ommon) N(oun) P(hrase) in Hote have the following order:

C(ommon) N(oun), M(odifier)<sub>1</sub>, M<sub>2</sub>, ..., Q(uantifier)/N(umeral), Dei(ctic) Form, Refl(exive)

Transcriptions in this paper are approximately phonetic, with the exception of the *ng* which is the digraph for the velar nasal and the *h* which is the velar fricative.<sup>3</sup> The phonology presents a number of problems, but is not within the scope of this paper (see Muzzey 1988).

## 3. POTENTIAL DETERMINER FORMS

In the following sections, details are given of the six forms in Hote that occur at the ends of noun phrases and are thus candidates for the status of 'determiner' in the language.

3.1 THE FORM *te*<sup>4</sup>

The form *te* occurs finally in a noun phrase when the head of the phrase is a singular noun:

- (1) ...*aleba i-vuyei*                      *valu adyang mavi te*.  
 until 3PL.REA-look/at stone hole good DET  
 ...until they saw a good cave. (Text 1.2)
- (2) *Ma ha-wa*                      *kising te hi*                      *aleba ha-dou*  
 then 3SG.REA-carry axe DET 3SG.REA/go until 3SG.REA-left  
*ha-ming*                      *kamung*.  
 3SG.REA-stand jungle  
 Then he carried an axe went until he left (it) it stands (in the) jungle. (Text 11.2)

<sup>2</sup>Negatives at this stage in the analysis are described as being composed of two words, *ming* and *ami*, both of which usually occur in a negative sentence, *ming* at the beginning and *ami* at the end, as indicated in the ordering.

<sup>3</sup>Some abbreviations used in glosses are:

1,2,3	first, second, third person
COMP	complementiser
DET	determiner (used initially for all six forms which are possibly determiners)
(FREE:...)	freer English translation (to capture semantics of a clause)
I, E	inclusive, exclusive
IRR, REA	irrealis, realis (differing forms of subject affixes, without any implication at this point that this is the correct analysis)
POS	possessive
POT	potential (prefix on verbs)
SG, DU, PL	singular, dual, plural

<sup>4</sup>This may be a reflex of the \**ta* 'indefinite article' reconstructed by Ross (1988:359) in PWO, but it would regularly reflect POC \**tai*.

- (3) *Menak te ha-mou kasukthom ba la-selo daim ba...*  
 bird DET 3SG.REA-live lake and 3SG.POS-neck long and...  
 A bird lives (at the) lake and its neck (is) long and... (Text 24.1)

The form *te* is also the most commonly used form for the numeral 'one':<sup>5</sup>

- (4) *baheng vi lahavu te* (Muzzey 1979:48)  
 hand(s) half including one  
 six
- (5) *lauming ba lahavu te* (Muzzey 1979:48))  
 ten and including one  
 eleven
- (6) *bung te* (Muzzey 1979:49)  
 whole one  
 twenty

In text, *te* occurs with the first occurrence of a noun phrase:

- (7) *...ma theilou e-lav unyak daim te ha-veng aleba...*  
 then 3PL 3PL.REA-build house tall DET 3SG.REA-go until  
 ...then they built a tall house going until... (Text 7.5)

In example (7), this is the first mention of this house in the text. But this does not imply that every first occurrence of a singular noun phrase in a text includes the form *te*:

- (8) *E-pengi ba yangaling 0 h-vyak ba...*  
 3PL.REA-fight and aeroplane 3SG.REA-put/down and...  
 They fought and (an) aeroplane came down and... (Text 5.4)

Let us more closely examine the occurrences of *te* in examples (7) and (8). In both cases, the nouns that the form *te* occurs with occur only once in the texts from which these examples are taken. However, the difference between the two noun phrases is in the question of whether the noun phrase continues to be a part of the following text or not. For *unyak* 'house' in example (7), the text continues with details of what happened in the building of the house. On the other hand, nothing further is either mentioned about *yangaling* 'aeroplane' or any aspect of it in relation to the text. Thus, though the form *te* seems to be a good candidate for a determiner that signals singular new information, its occurrence in the text is severely limited to nouns which are more 'prominent'<sup>6</sup> in some sense.

But, this is not the entire description of *te* with noun phrases, as can be seen in the following example:

- (9) *Ha-wa amena atu te ba ha-lei ba...*  
 3SG.REA-carry child DET DET and 3SG.REA-come/down and  
 He carried this child and he came down and...

In this clause, the form *atu* clearly marks *amena* 'child, children' as 'old information' (see section 3.2). What then does *te* indicate? A look at the context of this sentence helps to clarify the use of

<sup>5</sup>There is another form, *dongtom* 'one'; but the distribution of this form is as yet undetermined.

<sup>6</sup>I do not define 'prominent' in this paper. Note, however, that a noun which refers to an object which is never mentioned again, and about which no further predications are made in a text, would not be prominent.

both of these forms. A translation for this noun phrase that is actually closer semantically is the following: 'one of the aforementioned children'. Within the text, more than one child was in the cave and these are talked about before this sentence. However, at this point, only one of them is carried away. So, although the children as a group constitute old information, the one that is now being carried is newly mentioned and this child now achieves 'prominence' in the story.<sup>7</sup>

Though the most frequent occurrence of *te* is in numerals and in noun phrases, it also may occur by itself in a referring, but indefinite, sense, where it functions as the indefinite 'one' does in English:

- (10) *Te ha-naing abouleik te.*  
 one 3SG.REA-speak mouth one  
 One spoke one language. (Text 7.14)

How then can we characterise the form *te*? The following sums up what we have described above:

- (a) the numeral 'one' (in NPs);
- (b) 'singular prominent' specifier (in NPs);
- (c) indefinite referring form (occurring independently).

It is thus likely that *te* is a good candidate for determiner status.

### 3.2 THE FORM *atu*

A second form which frequently occurs with noun phrases is *atu*. This form occurs with a noun that has appeared previously in the text:

- (11) *...e-wa vak lo num ha-thak valu adyang mavi atu.*  
 3PL.REA-carry string bag and food 3SG.REA-go/up stone hole good DET  
 ...they carried the string bag and food it go up to this good cave. (Text 1.3)
- (12) *Valu atu adyang,<sup>8</sup> ha-bi-tak lov-yaing.*  
 stone DET hole 3SG.REA-push-up hole-another  
 This cave pushed up two places.  
 (i.e. The cave had two openings.) (Text 1.4)
- (13) *Anyou atu lo veng e-mou aleba daing.*  
 man DET and spouse 3PL.REA-stay until finish  
 This man and his wife stayed until finished. (Text 2.35)

Examples (11) and (12) provide further examples of the noun phrase *valu (adyang)* 'cave'. The first occurrence of the noun phrase was given in example (1) above with the form *te*. The man in example (13) has also been mentioned before in the text. The first time, *te* occurred with *anyou* 'man' and thereafter *atu* occurs with *anyou* in the story.

However, *atu* is not restricted to a previously mentioned noun phrase:

- (14) *...aleng yaha-yei ngambu being atu leik ha-louk...*  
 until 1SG.REA-see flood big DET now 3SG.REA-go/down  
 ...until I saw this big flood now go down... (Text 22.5)

<sup>7</sup>The child is mentioned after this sentence, and again when the experiences of the child are told to someone else.

<sup>8</sup>I don't have a good explanation at this point in the analysis for the occurrence of *atu* between *valu* 'stone' and *adyang* 'hole'.



In the text in which the example above occurs, this is the first occurrence of *ngambu* 'flood'. However, this flood is a historic event in the village. The speaker thus assumes that the hearers will know which flood is being talked about (and probably rightly so). We can say that this noun phrase is considered by the speaker to be within the knowledge of the hearers and is thus 'old information'.

Let us consider the following example:

- (15) *Ha-deng*                    *sebouk atu ba avi te lo veng louk...*  
 3SG.REA-walk(IRR) before DET and woman DET and spouse and  
 It walked before and a woman and her husband and...  
 (FREE: It was before and ...) (Text 1.1)

Here, *atu* occurs with the first phrase of the first clause of the text.<sup>9</sup> It occurs with the time word *sebouk* 'before' (or 'the time before'). This can also be considered 'old information', in the sense that past time is something within the knowledge of the hearer.

Added evidence that *atu* marks old information comes from the following example:

- (16) *Ha-pu*                    *theilou ha-mou bulivung atu:...*  
 3SG.REA-guide 3PL 3SG.REA-stay night DET  
 It guides them it stays this night:... (Text 16.44)

There is no previous indication of which night the speaker is referring to in this text. The speaker senses this and adds after this clause another one in which he clarifies which night he is talking about. Thus, a case in which *atu* is used to indicate 'old information', but in which the referent is perhaps not within the knowledge of the speaker requires further clarification.

Finally, we note that there are no number restrictions on the use of the form *atu*. Though in our examples thus far it has occurred with singular nouns, it may also occur with plural nouns:

- (17) *Ma avumalu tak atu leik e-n-eik...*  
 then people all DET now 3PL.REA-POT-sleep  
 Then all these people now slept... (Text 1.13)
- (18) *Ma bahengvi atu ming ev having ami e-le...*  
 then five DET NEG give belief NEG 3PL.REA-go/down  
 Then these five (who) didn't believe went down... (Text 16.63)

As might be expected of a form that marks known information, *atu* occurs independently in a referring role, where it refers to something already mentioned:

- (19) *E-langu*                    *atu ba...*  
 3PL.REA-hear DET and...  
 They heard this and... (Text 4.15)

In the above example, *atu* refers to something that had just been said. It also occurs with the indefinite noun *noum* 'something', to refer to previously given information:

<sup>9</sup>I initially call this a phrase, even though I have glossed the form *ha-deng* as '3SG.REA-walk', indicating that it is a verbal and this is actually some type of clause. However, the form is 'frozen' in a sense and means something like 'in' or 'during the time of', etc. Further analysis of verb phrases and clauses should throw light on the phrase or clause status of this construction.

- (20) *Orait*,<sup>10</sup> *ma noum atu alou na-yala Wapoum-being ma...*  
 OK then thing DET 1PL.IRR 1PL.IRR-know God-big then  
 Okay, then this thing we know (about) God then... (Text 16.123)

Here, the indefinite noun refers back to a particular thing that is known, since it has just been described in the text. Finally, we note that *atu* may occur with one of the other determiner-like forms:

- (21) *Ma e-n-dum aeinteik atu leik...*  
 then 3SG.IRR-POT-do DET DET now  
 Then he should do this now... (Text 16.62)

In this usage, *aeinteik* is a form which refers to the following clause (section 3.4), and *atu* marks that clause as old information in the text, even though it follows the forms. And indeed, the following clause is a repeat. These, then, are the features of the form, *atu*, relevant to this discussion:

- (a) marks old information (in NPs);
- (b) refers to previously mentioned nouns (occurring independently).

Again, *atu* is a good candidate for determiner status in Hote.

### 3.3 THE FORM *enteik* ~ *einteik* <sup>11</sup>

The form *einteik* is the one which most frequently occurs with time words:

- (22) *Avumena ining yeng yansing einteik eing.*  
 children 3PL.POS song afternoon DET EMPH  
 The children's singing (is) this afternoon. (Text 4.1)
- (23) *A-kopak ha-veng kamung kapo ha-toum wak*  
 1PL.REA-hide 3SG.REA-walk jungle inside 3SG.REA-become day  
*noumbeing einteik.*  
 many DET  
 We hid (walking?) inside (the) jungle for these many days. (Text 5.10)

In both examples, the time referred to in the text is the time of the occurrence of the event described in the text.

It also occurs if the place the speaker is in is talked about in the text:

- (24) *...aleba yaha-you long einteik.*  
 until 1SG.REA-arrive place DET  
 ...until I arrived this place (here). (Text 17.3)
- (25) *...ba e-veng kamung noumbeing einteik sapeng.*  
 and 3PL.REA-scatter jungle plenty DET all  
 ...and they scattered (in) all these plenty jungles. (Text 2.60)

In example (24), the speaker is referring to the village he is in right now. In example (25), he is referring to the jungles in which the village he is in is situated.

<sup>10</sup>This is a Tok Pisin word.

<sup>11</sup>Henceforth *einteik*.

*Einteik* can also specify other objects:

- (26) *Moulo alim tak einteik no-sak dongtom enga...*  
 2PL fish all DET 2PL.IRR-come/up one because  
 You, all these fish, come up one (place) because... (Text 24.8)
- (27) *...na teim ne-dav kasuk einteik thou.*  
 COMP time 3PL.IRR-dig lake DET away  
 ...that (in) time they will dig this lake out. (Text 24.9)

Even though the noun phrases in the above examples are not time or place, the form specifies that they are in the close vicinity of the speaker in the text.<sup>12</sup>

*Einteik* does not mark a noun phrase with regard to information structure and thus occurs with nouns which are regarded as old information and with nouns which are regarded as new information:

#### OLD INFORMATION

- (28) *Ma eiv nadokta einteik iom ba...*  
 then 3SG.REA/give little DET only and...  
 Then he gave this little only and... (Text 14.8)

#### NEW INFORMATION

- (29) *...ma uthuv einteik do.*  
 then possum DET plenty  
 ...then these possums (were) plenty.  
 (FREE: ...then there were plenty possums.) (Text 2.19)

In example (28), the little bit that this man gave to his wife has already been talked about. In example (29), there is no previous mention of possums and, in fact, these possums have just appeared in the vicinity of a trap.

In the one example in my material where *einteik* appears by itself, it refers to the place 'here':

- (30) *...ma thai e-leim einteik.*  
 then 3DU 3PL.REA-come/up here  
 ...then they two came up here. (Text 17.27)

A form that is apparently related to *einteik* indicates the time 'now':

- (31) *Ma mentek eing kam ka...*  
 then DET EMPH FUT so  
 Then now it will be like this, so... (Text 3.17)

Though *einteik* is generally used to refer to something that is close in space to the speaker, there is one example in which it is obvious that the object to which the noun phrase refers is not in the vicinity of the speaker:

- (32) *... "Ek, anyou einteik ha-sau ya ma..."*  
 eh man DET 3SG.REA-lie 1SG then  
 ... "Eh, this man lied to me then..." (Text 21.6)

<sup>12</sup>The speaker in this case is a bird that is trying to get some fish in one lake to move to another lake.

In this text the man (her husband) that the woman is speaking about has left to go hunting. Instead, the form here implies disgust with him (because of his lying). Just how this function of the form fits in with the other findings is not clear and I simply note it here.<sup>13</sup>

How then, is it best to characterise *einteik*? The following are two characteristics that can be seen from the observations above:

- (a) close to speaker (in NPs and occurring independently);
- (b) time of occurrence of event being described or now (in NPs and occurring independently).

Thus, the form *einteik* also is a good possibility for having determiner status.

### 3.4 THE FORM *aeinteik*

A form very similar to *einteik* is used to introduce both direct and indirect quotes:<sup>14</sup>

#### DIRECT QUOTES

- (33) *Ma yamalu ha-naing aeinteik eing, "Moulu..."*  
 then spouse 3SG.REA-say DET EMPH 2PL  
 Then her husband said this, "You...." (Text 1.6)
- (34) *Ma kiav ha-naing aeinteik, "E, ho-veng opali?"*  
 then kiav 3SG.REA-say DET eh 2SG.REA-walk who  
 Then (the) Kiav said this, "Eh, you walked (with) who?" (Text 5.17)

#### INDIRECT QUOTES

- (35) *...ha-naing ha-deng yai aeinteik nena batutek ma...*  
 3SG.REA-say 3SG.REA-walk(IRR) 1DU.E DET COMP now then  
 ...he said to us this that (it is) now then... (Text 17.18)
- (36) *... "O ya-leng-habi abou te aeinteik nena alalou..."*  
 oh 1SG.IRR-stomach-think talk DET DET COMP 1PL.I  
 ... "Oh, I think this talk that we..." (Text 19.10)

The quote following this form need not appear in the same clause:

- (37) *...enga anyou lokwangyu e-naing aeinteik.*  
 because man two 3SG.REA-talk DET  
 ...because (the) two men said this. (Text 24.8)

However, the clause in which *aeinteik* appears is repeated in the next sentence with the following quote:

<sup>13</sup>Another occurrence of this form is in equationals:

*Ma avi bahengvi einteik ma alou Kristen einteik*  
 then woman five DET then 1PL.I Christian DET  
 These five women (are the same as) we Christians. (Text 16.139)

However, until there is further analysis of such clauses, it is uncertain that this function differs from what I have described in this paper.

<sup>14</sup>I do not speculate in this paper as to the origin or function of the *a-* in *a-einteik*.

- (38) *Anyou lukwangyu e-naing aeinteik nena teim...*  
 man two 3PL.REA-say DET COMP time...  
 The two men said this that the time... (Text 24.9)

This kind of occurrence is probably a stylistic device used in discourse.

As shown above, *atu* (example (21), here repeated as example (39)) may occur with *aeinteik*:

- (39) *Ma e-n-dum aeinteik atu leik h-eiv theilou leik...*  
 then 3SG.IRR-POT-do DET DET now 3SG.REA-give 3PL now  
 Then he would do this (aforementioned) now: he gives them now... (Text 16.62)

In example (39), there is no indication of a quote being present in the sentence. Instead *aeinteik* refers to the next clause, which has no other quotation markers.<sup>15</sup>

Finally, note that generally *aeinteik* does not occur with a noun (examples (33), (34), (37) and (38)). In example (35), it seems unlikely that *aeinteik* is in the same noun phrase as the pronoun *yai* 'first person dual exclusive' since pronouns generally have no following deictic form. Example (36) needs some explanation as *aeinteik* seems to specify *abou* 'talk':

...*abou te aeinteik...*  
 talk DET DET  
 this talk

However, since there is already a determiner that occurs with *abou* (*te* DET) it seems more reasonable to say that *aeinteik* is referring to the following indirect quote in line with other occurrences of *aeinteik*.

Thus, to characterise *aeinteik*, the following points are reasonable:

- (a) refers ahead to direct and indirect quotes;
- (b) possibly refers to other following clauses, the characteristics of which are yet to be described.

Since there are no occurrences of this form where we can definitely say that it occurs with noun phrases, it is not a likely form for determiner status. Instead it resembles an independent demonstrative form.

### 3.5 THE FORM *intu*

The form *intu* neither specifies noun phrases in any identifiable way nor does it refer to following clauses as such. Instead if we are to attempt to identify any referring function, in some cases it can be interpreted as referring to an 'event'. (In this case, I use the term 'event' to indicate an actual happening that was described, not something said.)

- (40) *Youv intu ma theilou e-lav unyak daim te...*  
 finished DET and 3PL 3PL.REA-build house tall DET  
 Finished this then they built a tall house... (Text 7.5)

<sup>15</sup>It is true that further analysis may show the following clause to be a type of indirect quote. However, that would not change the basic analysis of this form.



- (41) ...*ba intu alou da na-pesang...*  
 and DET 1PL.I REFL 1PL.IRR-prepare  
 ...and thus, in that way we prepare ourselves... (Text 16.132)

In example (40), *intu* refers to a described migration from Australia. In (41), it refers to some preparations that were described. In both of these examples, note that *intu* would be interpreted as the head of a phrase.

However, there are cases in which the form seems to occur with noun phrases:

- (42) *Ha-yuv lavuwak intu ma...*  
 3SG.REA-blow conch/shell DET then  
 He blew the conch shell then... (Text 23.34)
- (43) ...*nenā wak intu ma thai ne-pum i enduk bu intu.*  
 COMP day DET then 3DU 3PL.IRR-meet 3PL at place DET  
 ...that this day then they two should them at this place (both designated). (Text 23.14)

In neither case does there seem to be any identifiable specifying function. In particular, in example (42), the best translation continues to be something like 'thus'. This is not as good an interpretation for the occurrences in example (43), however.

But there are other even more problematic occurrences:

- (44) *Yani intu ha-naing na teim yani...*  
 3SG DET 3SG.REA-say COMP time 3SG  
 He said that he... (Text 16.105)
- (45) *Intu yani anyou ha-wa avi lukmuk...*  
 DET 3SG man 3SG.REA-carry woman new  
 He (is like) a man he gets a new woman... (Text 16.127)
- (46) *Ma lokbuk intu eing ma ha-viyou ha-naing...*  
 then morning DET EMPH then 3SG.REA-get/up 3SG.REA-say  
 Then (it is) morning then he got up he said... (Text 11.3)

Looking at examples (45) and (46), we see the possibility that *intu* may be verbal in some way. Let us explore this possibility. First, we note that there is a possible verbal morphological division for this form:

*i-n-tu*  
 3SG.IRR-POT-become, be/like

All these morphemes are fully attested as commonly occurring in verbal constructions in the language. The following is an example of the verbal root, *-tu* 'to become':

- (47) *Yeilou na-tak anyou te ka i-tu*  
 1PL.E 1PL.IRR-mark/allow man DET so 3SG.REA-become  
*yeilou-ning anyou being.*  
 1PL.E-POS man big  
 We mark a man so he becomes our leader. (Muzzey 1989)

In example (47), note that the bold form is a verbal for the result clause and is similar in form to *intu*. The addition of the potential morpheme *-n* could account for the difference between the occurrence in this example and the other occurrences.<sup>16</sup>

If the *intu* we are describing is a form of the verbal, *tu*, 'become' then we would expect some semantic support for this analysis in the glosses. And in fact, in every instance, it is possible to give a meaning for this form that is something like the following: 'has come into being and is, became'. Example (46), for instance, could be retranslated as 'Then morning (be)came then he got up he said...' In (45), *intu* could be considered an introductory verbal: 'It happened, he...' And likewise in the other examples, *intu* lends itself more readily to being a verbal of the type I have described.

And, in fact, we find that not only does *intu* occur where it is reasonable to analyse it as a type of verbal in a clause and in serial verb constructions, but there is at least one example where it can be analysed as the last verb in a serial verb construction:

- (48) *Ma yani ha-you intu.*  
 then 3SG.REA 3SG.REA-arrive 3SG.REA-POT-is  
 Then she arrived. (Text 4.4)

If we do not analyse this as a verbal with the aspect of 'having come into being', we are left with the problem of identifying it as the head of a noun phrase and assigning it semantics which 'fit' with other occurrences of the form, a task, which at this point, does not seem feasible. If my analysis of *intu* as being verbal is correct, and there is no identifiable specifying function for it when it seems to occur with nouns, then this form is not a candidate for determiner status.

### 3.6 THE FORM *eing*

Because *eing* occurs finally in noun phrases, it seems to be a candidate for determiner status:

- (49) *Ma avi atu ba h-eiv yeng ek avumena eing.*  
 then woman DEI and 3SG.REA-give song for children DET  
 Then this woman and she gave songs for children. (Text 4.3)
- (50) *Malak lu eing e-mou Butamum.*  
 village three DET 3PL.REA-stay Butamum  
 The three villages are at Butamum. (Text 23.2)
- (51) *Ma menak eing ha-mou aleba bok lokaing oyang ba mi.*  
 then bird DET 3SG.REA-stay until animal bone nothing and NEG  
 Then the bird stayed until bone nothing and not. (Text 24.3)

However, it also occurs at the end of verbal phrases:

- (52) *Ba leik youv eing ma ming ho-thak...*  
 and now finished DET then NEG 2SG.REA-come/up  
 And now finished then you don't come up... (Text 25.31)

<sup>16</sup>It is admittedly hard to explain why a 'potential' form should occur with something that is accomplished. However, my suspicions are that the 'potential' form needs further semantic analysis. And, of course, this *n-* might be something else entirely.

- (53) ...*ma teim ming ha-toum alalou na-tak na-you*  
 then time NEG 3SG.REA-be/able 1PL.E 1PL.IRR-come/up 1PL.IRR-arrive  
*malak leng paling ami ma mi eing.*  
 house sky really NEG then NEG DET  
 ...then at this time, it didn't allow us to reach the sky home really then not. (Text 16.95)
- (54) ..."*Alai a-veng leik eing alai na-ng-geik kamung.*"  
 1DU.I 1PL.REA-walk now DET 1DU.I 1PL.IRR-POT-sleep jungle  
 ..."We two walk now we sleep (in the) jungle". (Text 21.2)<sup>17</sup>

In some cases, *eing* occurs with a time word:

- (55) *Leik eing Wapoum-being h-eiv about bu te...*  
 now EMPH Good-big 3SG.REA-give talk place DET  
 Now God gave talk (at) a place... (Text 16.64)

It may also modify a negative:

- (56) ...*alou-ning auk ming (a)eing*<sup>18</sup>  
 1PL.I-POS knowledge NEG EMPH  
 ...our knowledge (is) not. (Text 16.52)

*Eing* may also occur after *einteik*:

- (57) ...*ma yani ha-naing hadeng i einteik eing...*  
 then 3SG 3SG.REA-say to 3SG.OBJ DET EMPH  
 ...then she said to him this... (Text 4.12)
- (58) *Avumena ining yeng yansing einteik eing.*  
 children 3PL.POS song afternoon DET EMPH  
 (It is) the children's songs this afternoon. (Text 4.1)

and before and after *intu*<sup>19</sup>

- (59) *Ma loukbouk intu eing ma ha-viyou ha-naing...*  
 then morning DET DET then 3SG.REA-get/up 3SG.REA-say  
 Then, the next morning then he got up he said... (Text 11.3)
- (60) *Kupik eing intu theilou e-wa ba...*  
 body DET DET 3PL 3PL.REA-carry and  
 The body they carry and... (Text 16.84)

Frequently, *eing* occurs in the closing line of a text:

- (61) *Leik youv eing.*  
 now finish DET  
 Now (it is) finished. (Text 6.11)

<sup>17</sup>It may be that example (54) is actually an instance of *eing* occurring at the beginning of a clause.

<sup>18</sup>An unanswered question is whether this form with initial *a* on is a different form or not. For the purposes of this paper, I have found no difference in either semantics or occurrence and so assume that they are the same (see fn. 14).

<sup>19</sup>I continue to gloss this form as DET, so as to identify it as a form discussed in this paper.

The most frequent semantics given for the form is that of 'emphasis'. Because of the relatively free occurrence of *eing* and its frequent interpretation as simply being an emphatic form, I suggest that this form is adverbial, rather than a determiner.

#### 4. DETERMINERS

In this paper, I have given evidence for eliminating three forms from possible determiner status:

<i>aeinteik</i>	'independent demonstrative'
<i>intu</i>	'verbal, become'
<i>eing</i>	'emphatic adverb'

Thus, three forms are likely determiners in Hote:

<i>te</i>	(a) numeral for 'one' (in NPs); (b) singular, prominent new information (in NPs); (c) indefinite referring (occurring independently).
<i>atu</i>	(a) old information (in NPs); (b) refers to old information (occurring independently).
<i>einteik</i>	(a) close to speaker (in NPs and occurring independently); (b) time of event and now (in NPs and occurring independently).

All three of these forms have a dual function. They occur both with noun phrases and independently. We can say, however, that they have some kind of determiner status, since they occur most frequently with noun phrases. However, it is obvious that they are not exclusively determiners (or forms that specify noun phrases). What are the theoretical considerations of forms that occur as these do?

#### 5. THEORETICAL CONSIDERATIONS

I would like to suggest that one of the most important implications of the analysis of determiner-like forms in this paper is the following:

It is possible that the semantics of these forms is the key factor in their occurrence in noun phrases.

Why is this true? First, it needs to be noted that it is common in Hote to have no determiner at all with a noun phrase:

- (62) ...*lo e-wa*                      *vak*                      **0** *lo num hathak...*  
       and 3PL.REA-carry string bag **0** and food inside  
       ...and they carried (the) string bag and (the?) food inside... (Text 1:3)

Thus, the concept of definiteness is not a prime factor in the occurrence of these forms. And, as we saw above, the features of old/new information, prominence in text, and closeness to speaker are the semantic features relevant to the occurrence of these forms. Note that the information and prominence features are actually features dependent on context, while the closeness to speaker is 'outside' context. I would thus like to suggest that, in Hote, at least, these forms are not particular

forms that only specify noun phrases (as we might think of determiners), but rather forms that relate the noun phrases to context, both textual and 'outside' context. With such a consideration, context (or discourse) becomes an important factor in the description of forms that appear in phrases and independently (as possible heads of phrases).

#### BIBLIOGRAPHY

- BRADSHAW, Joel**, 1978, Notes on subgrouping in the Huon Gulf area. *WPLUH* 10/1:49-83.
- HOOLEY, Bruce A.**, 1971, Austronesian languages of the Morobe District, Papua New Guinea. *Oceanic Linguistics* 10:79-151.
- MUZZEY, Marguerite**, 1979, Hote grammar essentials. Manuscript, Summer Institute of Linguistics, Ukarumpa, Papua New Guinea.
- 1988, Hote phonology essentials. Manuscript, Summer Institute of Linguistics, Ukarumpa, Papua New Guinea.
- 1989, Computer printout of Hote wordlist. Summer Institute of Linguistics, Ukarumpa, Papua New Guinea.
- ROSS, M.D.**, 1988, Proto Oceanic and the Austronesian languages of Western Melanesia. *PL*, C-98.



## CONSONANT DISSIMILATION IN MAORI

RAY HARLOW

### 1. INTRODUCTION

Visible from much of the eastern Bay of Plenty in the North Island of New Zealand is an island, which lies just off the city of Whakatane. It reminds one of a whale, at least at a distance, and in recognition of this bears the name *Moutohora*, of which the English name 'Whale Island' is a translation. The name is derived from *tohora* 'whale' and *mou*, a form of the word *motu* 'island' which otherwise occurs to the best of my knowledge, only in another island name, *Moutoki*, and in the compound *moutere* 'island' (according to informants and to Williams (1971), the form *motutere* also occurs). The temptation to see the form *mou-* in all three of these words as the result of a complete dissimilation of the *t* in *motu* is of course great, and this paper is an attempt to explore the implications of the postulation of such a process for Maori. In the course of this exploration, I will make some proposals about the process itself, but also about some otherwise puzzling aspects of Maori morphology.<sup>1</sup>

Another word which shows a similar variation is the 'tribal prefix' *ngāi* ~ *ngāti*, which occurs in very many tribal names, and means originally something like 'the descendants, the clan (of)'. It is striking that, with one exception, the form *ngāi* occurs only before names beginning with *t*, for example, *Ngāitahu*, *Ngāitai*, *Ngāi Te Rangī*, and so on. The same prefix appears in the expressions *ngāi tāua* 'the-clan-of us.INCL.DU' and *ngāi tātou* 'the-clan-of us.INCL.PL', used by Maori to refer to the race as a collective.

There are two kinds of counterexample to the conclusion that here as well dissimilation has led to the allomorphy; however, an account can be given of these which does not necessitate the abandonment of the proposed analysis.

Firstly, there are many tribal names where the form *Ngāti* is followed by *t*, for example, *Ngāti Tūwharetoa*, *Ngāti Tai*, *Ngāti Toa*. Secondly, there is the expression *ngāi māua* 'the-clan-of us.EXCL.DU' (Williams 1971: *ngāi*) alongside the 'expected' *ngāti māua*. An explanation of these counterexamples which is consistent with the allomorphy's having arisen through a phonological process would point to the fact that the process was never more than optional (cf. *moutere* ~ *motutere* above), and to the close connection between the two sets of expressions involving 1st person non-singular pronouns. The tribal names in *T-* which nonetheless have *Ngāti* as the prefix are simply exceptions to an optional process, and the form *ngāi māua* is analogical.

<sup>1</sup>Many thanks to Jeffrey Waite, who read an earlier version and made very helpful comments.

Similar remarks can be made about another morpheme which shows comparable variation of form. The Proto Polynesian prefix \*faka- 'causative' appears in Maori in two forms: *whaka-*, far and away the more frequent form, and a variant, *whā-*. Some examples: *whakainu*, *whāinu* 'to give to drink', *whakakoekoe*, *whākoekoe* 'tickle'. I propose that this variation is also ultimately the result of the same process of consonant dissimilation, and that the counterexamples (see below) are the result of the process's being optional and of analogical extension of the domain of the originally phonologically conditioned allomorph.

Precisely similar alternations in the form of this prefix occur in related languages, though there they are restricted to the originally conditioning environments. For example Fijian: *vaka-* > *vā-* before velars, cf. Arms (1989); Hawaiian: *ho'o-* > *hō-* before a glottal stop (< *k*), Elbert and Pukui (1979:76).<sup>2</sup> However, in Maori, not only does the form *whaka-* also occur before *k*, the form *whā-* occurs before sounds other than *k*.

Since the process being invoked here was optional, doublets were produced in precisely similar environments. This outcome was exploited by Maori on occasion in order to disambiguate homophones. One sometimes finds homophonous roots selecting different forms of the causative prefix, for example, *whākuru* 'pelt' < *kuru* 'pelt' but *whakakuru* 'weary' < *kuru* 'weary', *whāngongo* 'administer liquid food to someone' < *ngongo* 'suck' but *whakangongo* 'neglect, pay no attention' < *ngongo* 'low born'. If the process leading to the allomorphy was 'opaque' in this way, there was nothing to prevent the analogical application of *whā-* to other contexts. Appendix 1 contains all the *whā-* forms I have been able to find.

What these three alternations have in common can be formulated as the following optional process:  $C_1V\#C_1V \rightarrow V\#C_1V$ , where # is some kind of morphological boundary,  $C_1$  is *k* or *t*, and the deleted *C* is not word initial. In all the cases looked at so far, the dissimilated *C* is at least three moræ<sup>3</sup> from the end of the word.<sup>4</sup>

## 2. PASSIVES

Granting the existence of such a process, and leaving aside the question of its status as a rule, an historical change, etc., it is of interest to see whether it plays a role in other alternations as well. Since Maori phonotactics do not admit consonant clusters, the effect of the dissimilation process is always to bring together two vowels otherwise separated by a single consonant. If these are two like vowels, they become a single long vowel by the process of syllabification,<sup>5</sup> as in *whaka-* versus *whā-* above.

One group of forms which are puzzling but which gain an explanation if viewed as instances of the dissimilation process consists of the passives of reduplicated verbs. A frequent pattern is exemplified by the pair: *pupuhi* ~ *pūhia* 'to blow or shoot, to be blown or shot', and a full list of the verbs which follow the pattern is given in Appendix 2. Accommodating this pattern under the proposed consonant dissimilation entails:

<sup>2</sup>In fact, the situation in Hawaiian is more complicated, cf. Elbert and Pukui.

<sup>3</sup>For the applicability of the notion *mora* to Maori phonology, cf. Bauer (1981).

<sup>4</sup>A few other isolated examples are suggestive: *matatau* ~ *mātau* 'know', *matatangata* ~ *mātangata* 'a univalve mollusc', *pararaki* ~ *pāraki* 'land wind'.

<sup>5</sup>Many of the examples adduced below raise interesting questions about the correct account of Maori syllabification. Suffice it to mention here that these examples support the analysis of surface long vowels as (underlyingly) a sequence of two like vowels.

(a) accepting that the process applies to consonants other than just *t* and *k*;

(b) accepting that the process is mirror-image, that is, that  $C_1V\#C_1V \rightarrow C_1V\#V$  also occurs (note that the distance of the dissimilant is still at least three moræ from the end, indeed, that is what prevents dissimilation in the active form);

(c) accepting that the passive of *pupuhi* is underlyingly *pupuhia*;

(d) accepting that there is some sort of boundary following a reduplication syllable.

As is well known, universally it is liquids and nasals which tend to provide the majority of examples of dissimilation, both partial, as in Spanish *hombre* < Latin *hominem*, or complete, as in French *faible* < Latin *fleibilem*. Thus, if in a language *k* and *t* undergo some dissimilation process, it is not only to be expected that *p* will also undergo it, but also that the more 'sonorous' classes of consonant will undergo it as well, and if possible to a higher degree. Again, from a universal perspective, bidirectionality of dissimilation processes should not surprise, cf. the examples already quoted.

Passive verbs in Maori are generally formed by the addition of a suffix of the general shape *-Cia*,<sup>6</sup> a very frequent allomorph of which is *-a*, for example, *patu* ~ *patua* 'strike', *horoi* ~ *horoi-a* 'wash'. This and the other allomorphs of *-Cia* show some preference as to the shape of stem to which they are appended, however, it is easy to find examples of the use of *-a* suffixed to stems consisting of three unimoraic syllables, for example, *takahi* ~ *takahia* 'tread'. There is thus nothing in the way of postulating the underlying form *\*pupuhia*,<sup>7</sup> etc. for the passives in Appendix 2.

The postulation of a boundary<sup>8</sup> of some kind cannot be directly justified in any non-circular way, since I am not aware of any process other than the dissimilation under discussion which is manifest in partially reduplicated forms. However, two points should be made in this connection.

(a) With the possible exception of two dubious cases to be mentioned below, I know of no instance of dissimilation within a stem. Thus, if the proposed analysis is correct, there is some kind of difference between the  $C_1VC_1$  sequence as part of a stem, and a  $C_1VC_1$  sequence in partial reduplication.

(b) There is some phonological indication that **complete** reduplication involves the insertion of boundaries. We will see later some evidence which suggests that at least some diphthongs in Maori are underlyingly disyllabic. If this is so, then there is a process which combines, for instance / $\$a\$/$ , to a monosyllabic diphthong / $\eta ai$ /. This process is inhibited by some sort of boundary in exactly the cases where the *a* and the *i* are brought together by complete reduplication. Thus, *pāinaina* 'warm oneself, bask', with reduplication of the component *ina* is superficially pentasyllabic. Precisely the same effect is found in words derived by prefixation of *whaka-* to words beginning in *i-*, for example, *whakairi* 'hang up (vtr)', which is tetrasyllabic.

The two possible exceptions to the requirement that a boundary should intervene between the two consonants involved in the dissimilation are:

<sup>6</sup>This is the conventional abbreviation for the paradigm in Polynesian linguistics.

<sup>7</sup>I use the symbol \* with systematic ambiguity to mean either 'reconstructed' as in 'Proto Polynesian *\*faka-*', cf. above, or simply 'unattested' as here.

<sup>8</sup>An alternative formulation of this is to claim (following Marantz's 1982 discussion of the interaction of reduplication and phonological rules) that our consonant deletion process is 'cyclic', a rule which applies only in derived contexts, but not entirely within a morpheme.

(a) *kuku* 'nip', which has two passive forms: *kukua*, *kūngia*. If the latter form is from *\*kukungia* by the loss of the medial *k* through the process we are discussing, then either the process can occur within a stem, that is, the initial *k* has dissimilated the second, or the dissimilation has been conditioned by the /ŋ/ of the suffix. Either event would require modification to the formulation of the process hazarded above.

(b) *kūtai*, 'mussel' in some dialects, may be a compound of *kuku* 'mussel' + *tai* 'sea'. If this etymology is correct and if the syllable *kū* from *kuku* exemplifies our dissimilation process, then this is a counterexample to the formulation requiring the presence of a boundary.

Neither of these cases being entirely clearcut, we will maintain the formulation given above for now.

There are two sorts of exception to the pattern: *pupuhi* ~ *pūhia*. In one kind, reduplicated stems have no reduplication in the passive form and no long vowel either. Examples of this include: *tītiro* ~ *tīrohia* 'look at', *momotu* ~ *motuhia* 'sever', *rarahū* ~ *rahua* 'seize', *tataū* ~ *tavia* 'push (a sliding board, as door)'. Clearly, in cases of this kind, the passive is simply derived from the simplex underlying the active, rather than from the active form itself.

In cases of the other sort, non-reduplicated actives have passive forms with long vowels, such as *riri* ~ *rīria* 'be angry (with)', *whai* ~ *whāia* 'follow', *whakarere* ~ *whakarērea* 'abandon'. Again, this would not seem an insurmountable problem for the approach ventured here. Such cases could be seen either as instances of simplex active versus reduplicated (plus dissimilated) passives or as instances of an analogical extension of an originally phonologically motivated long vowel pattern in passive forms.

Before leaving passive formation, there are three further active-passive pairs which deserve comment here. All three are quite idiosyncratic, yet may receive an explanation if viewed in the present context. They are: *ārahi* ~ *arahina* 'lead', *tiki* ~ *tīkina* 'go and get', and *noho* ~ *nōhia* (also *nohoia*) 'sit, inhabit'. The last of these is in my view the most straightforward, the passive deriving from *\*noho-hia* by loss of the first *h*. The deverbal noun *nōhanga* can be derived in the same way from *\*noho-hanga*. The alternative passive and nominalised forms (*nohoanga*) cannot be derived in a similar way from the same underlying forms without sacrificing the otherwise general requirement that the lost consonant is at least three moræ from the end of the word. This is no great difficulty, however, since the suffixes *-ia* and *-anga* as allomorphs of *-Cia* and *-Canga*, respectively, are widely attested. A number of explanations are available for the other two pairs of forms, and I am not aware of a way of choosing between them. The active *ārahi* with its odd long vowel may be from *a-arahi*, that is, a partial reduplication, the passive (and nominalisation, *arahanga*<sup>9</sup>) being from the simplex. More adventurously perhaps, the active may be a compound of *ara* 'way' and *rahi* 'big', and the passive simply *ara-hina*, *hina* being an allomorph of *-Cia*. That is, the syllables *-hi-* which occur in both active and passive forms may not be cognate! Another of the words for 'lead' is of this form, *ara-taki*, and two of the other words for 'way' are compounds involving adjectives meaning 'big', *hua-nui* (*nui* 'big') and *hua-rahi* (*rahi* 'big'). Finally, *tīkina*. Three accounts spring to mind.

(a) < *\*ti-tiki-na*, i.e. passive is from the reduplicated form,

(b) the vowel is long by analogy, cf. above on *rīria* etc.,

(c) < *\*tiki-kina*, this suffix being yet another allomorph of *-Cia*.

<sup>9</sup>*ārahi* ~ *arahanga* belongs to a small set of verbs whose stem ends in an *-i* which is lost before the nominalising suffix. On these, cf. Williams (1971:xxxvii) and Krupa (1966:54).

### 3. REDUPLICATION

Another major area of Maori phonology where alternations of long and short vowels occur is in reduplication. The forms and functions of reduplication in Maori constitute a radically under-researched area.<sup>10</sup> This paper will not alter that situation to any great extent, but will only point to some formal features of reduplication which are superficially puzzling, and make some proposals about how to account for them.

In this discussion I shall be concentrating on the reduplication of trimoraic stems. The most common type of lexical stem in Maori is bimoraic, and undergoes both partial (initial underlying syllable) and complete (both underlying syllables) reduplication. Thus, from *paki* 'slap' come both *papaki* and *pakipaki*. There do exist however a very considerable number of trimoraic stems, which show a great variety of types of reduplication.

One of the puzzling aspects of the reduplication of longer stems in Maori is the existence of some doublets. Thus, corresponding to the simplex *pātai* 'ask', there exist both *pātaimai* and *patapatai* as frequentative forms. Krupa (1966) was aware of this and comments:

A variety of poly-vocalic words may be subdivided in two ways: (1) as consisting of a reduplicated root morpheme and a prefix, or (2) as consisting of a reduplicated root morpheme and a suffix. For instance *hokai* [sic, Williams 1971 has *hōkai*, RBH] 'extended, far apart' may be regarded as composed of the prefix *hoo-* ~ *ho-* and the root *kai* (because *hoo-kai-kai*<sup>11</sup> 'extend and retract alternately' occurs), or as composed of the root *hoka* and the suffix *-i* (because *hokahokai* 'extend' occurs as well). (1966:33)

He explains the *pātai* set in the same way.

Whether or not it is correct that a language can on a regular basis assign competing analyses to words, it is not in fact necessary to postulate a complex structure for stems of this type. Just as the more common bimoraic stems have more than one pattern of reduplication, partial and complete, available for their expansion, so too we can recognise the existence of a variety of patterns for trimoraic stems<sup>12</sup> without having to reduce these to extensions of the bimoraic type.

In particular, I am proposing that there are **four** patterns of reduplication available to trimoraic stems, which with the simplex makes **five** forms in which such stems may appear. I know of no stem which appears in all five, and only one where four are attested. Further, many stems appear only in one or more of the reduplicated forms and there is a gap in the simplex slot. Which of these forms any particular stem assumes is a matter for the lexicon, though, as with reduplication generally, research may reveal the existence of regularities which have escaped me.

Some of the reduplication patterns are rendered slightly opaque by the (in two cases, regular, in others, sporadic) operation of our consonant dissimilation process. Indeed, it is precisely because I believe that this process plays a role in these forms that I am mentioning them at all. Conversely, the postulation of the process of conditioned consonant deletion 'explains' some of the odd features of these patterns.

<sup>10</sup>Krupa (1966) deals with formal aspects of this phenomenon within his general treatment of word formation.

<sup>11</sup>Krupa adds the footnote: 'Only root morphemes may be reduplicated in Maori. (Unlike the affixes and particles).'

<sup>12</sup>No doubt many of these are etymologically complex, though it is questionable how far this is true synchronically.



The account I am giving here makes use of the approach known as Moraic Phonology,<sup>13</sup> and claims that the underlying syllable structure of Maori is  $\sigma = (C)V$ ,<sup>14</sup> and all  $\sigma$ 's are uniformly of one mora ( $\mu$ ) length, the mora inhering in the peak. This is the level at which reduplication rules, among others, operate. Subsequent rules combine some sequences of vowels into phonetically long vowels and diphthongs,<sup>15</sup> and thus surface syllables of length greater than one  $\mu$ .

Given a basic trimoraic stem shape of  $\sigma_1\sigma_2\sigma_3$ , the four reduplication patterns (with a couple of examples each for the meantime) are:<sup>16</sup>

1. reduplication of the first syllable:  $\sigma_1\sigma_1\sigma_2\sigma_3$ :  
*hohoata* < *hoata* both meaning 'the moon on the third day, pale, colourless';  
*a-anini* < *anini* both meaning 'giddy, aching (of the head)'.
2. reduplication of the first two syllables:  $\sigma_1\sigma_1\sigma_2\sigma_2\sigma_3$  (with dissimilation of the repeated consonant in  $\sigma_1\sigma_1$ ):  
*tāweweke* 'slow, dilatory' < *taweke* 'linger';  
*mānenei* < *manei* 'reach out to'
3. reduplication of the first two syllables:  $\sigma_1\sigma_2\sigma_1\sigma_2\sigma_3$ :  
*takatakai* 'wind round and round' < *takai* 'wrap up';  
*riariaki* < *riaki* 'raise'
4. reduplication of all three syllables:  $\sigma_1\sigma_1\sigma_2\sigma_3\sigma_2\sigma_3$  (with dissimilation of the repeated consonant in  $\sigma_1\sigma_1$ ):  
*pākarukaru* 'break in pieces (vtr)' < *pakaru* 'broken';  
*pāhūhū* 'pop, crackle' < *pahū* 'explode', i.e. *pa-a-hu-u-hu-u* < *pa-hu-u*.

Of these, Pattern 4 is very much the most frequently attested, as even a casual look at Williams (1971) will reveal. Pattern 3 is also fairly common, while Pattern 2 is rather rare. Appendix 3 is a table of simplex and reduplicated trimoraic stems, with no claim that the listings are exhaustive. Indeed, if any of these patterns are in any sense productive, exhaustive listings are impossible.

There are four places in the patterns stated above where the complete dissimilation of consonants occurs. In all cases, the provisional formulation arrived at above describes the site of deletion and the relationship to the conditioning consonant.

(a) Pattern 1: In some cases, the  $C_1VC_1V$  of  $\sigma_1\sigma_1$  is dissimilated to CVV, thus: *tākai* 'bandage' < *takai* 'wrap up' (i.e. *\*ta-takai*); *pākaru* 'break in pieces (vtr)' < *pakaru* 'broken'. In one case, there exists a pair of forms in one of which this has occurred and in the other of which the consonant is preserved: *papatū* 'screen for defensive purposes' ~ *pātū* 'screen, wall' < *\*patū*.

(b) Pattern 2: The  $C_1VC_1V$  of  $\sigma_1\sigma_1$  is always dissimilated to CVV, however, the  $C_2VC_2V$  of  $\sigma_2\sigma_2$  is never dissimilated to CVV. Thus, the pattern is always as given above.

<sup>13</sup>Cf. for example Hayes and Abad (1989 and ref. there). This approach is anticipated for Maori by Bauer (1981).

<sup>14</sup>Statement of these patterns in terms of 'syllables' is for convenience only, and should not be taken to imply that these patterns must be stated in these terms. Indeed, a translation into Marantz's 1982 C-V skeleta plus copying of (parts of) the morpheme is very simple and loses nothing.

<sup>15</sup>This aspect of Maori phonology must be left for another occasion, though a number of phenomena mentioned here provide evidence for such an account.

<sup>16</sup>Note by the way, that at least some of these patterns are inherited: cf. Fijian (see Arms 1989) *balabalavu* 'very long' < *balavu* 'long', *butobutō* < *butō* 'dark', Pattern 3; *kakadresudresu* 'torn to shreds' < *kadresu* 'torn', Pattern 4. Similarly, for Hawaiian examples, cf. Elbert and Pukui (1979:66).



(c) Pattern 4: the  $C_1VC_1V$  of  $\sigma_1\sigma_1$  is always dissimilated to CVV.

(d) Pattern 4: if  $\sigma_2 = \sigma_3$ , then the first occurrence of  $C_1VC_1V$  of  $\sigma_2\sigma_3$  may be dissimilated to CVV. This is in keeping with the stipulation that the deletion site must be at least three moræ from the end of the word. Thus, *pōwhīwhīwhi* = *pōwhīwhīwhīwhi* < *powhiwhi* all meaning 'tangled'.

Among the sets given in Appendix 3 is one representative of a phenomenon which should be mentioned here though it bears only indirectly on the primary object of this discussion. The three obviously related forms *mōnehu* 'die, expire', *mōnenehu* and *mōnehunehu* 'indistinct' are listed as the Patterns 1, 2 and 4 reduplications of an unattested \**monehu*. There exist many other Maori words of the form  $C_1V_1V_1C_2V_2C_3V_3$  (Pattern 1), for which there exists neither a  $C_1V_1C_2V_2C_3V_3$  ('simplex') form nor a reduplication of the shape  $C_1V_1C_2V_2C_1V_1C_2V_2C_3V_3$  (Pattern 3). That is, there exists no surface form with an initial short syllable.

Many of these words have a Pattern 4 reduplication, thus for example, *hāmama* 'open, gaping' ~ *hāmamamama* 'yawn', or *tūpeke* 'jump' ~ *tūpekepeke* 'jump about'. Pattern 2 is much rarer, but is exemplified above with *mōnenehu*. These words can of course be derived as implied in Appendix 3; that is, with underlying simplexes of the form  $C_1V_1C_2V_2C_3V_3$ , which chance not to surface or to have Pattern 3 reduplication. An alternative (and, in many instances, more likely) analysis would take  $C_1V_1V_1C_2V_2C_3V_3$ , that is, *mōnehu*, *tūpeke*, *hāmama*, to be the simplexes, and allow a reduplication pattern which consisted in a copy of the last two moræ only. Such a pattern is to be found with some words which are unequivocally tetramoraic in their simplex form, e.g. *porotiti* = *porotititi* 'disk, revolving', *karawhai* = *karawhaiwhai* 'enclose in a net', *raunui* = *raununui* 'broad'. Many of these words are clearly derived in the first place by prefixation to a stem which, in many cases, exists as a free form. Thus, besides *tūpeke* 'jump', there is also *peke* 'jump'; besides *raunui* 'broad', there is *nui* 'big'.

#### 4. CONCLUSION

What I have tried to do in this paper is to demonstrate the existence of an optional process in Maori which deletes one of two consecutive like consonants, so long as the deleted consonant is neither word initial nor less than three moræ from the end of a word. The process is a cyclic rule in the technical sense used by Marantz (1982) in his discussion of reduplication, and by others.

The process provides an account of some instances of allomorphy in Maori and of some of the details of reduplicated forms. Reduplication is clearly a complicated and under-researched area in Maori, with respect both to its formal properties and to the meanings expressed by it. Similarly under-researched is Maori phonology as a whole. This paper has brought to light some problems in these areas, and suggested starting points for further work.

#### APPENDIX 1

##### *whā-* ~ *whaka-*

*whāinu* = *whakainu* 'give to drink'

*whākana* 'make grimaces' < *kana* 'stare wildly'

*whākanakana* 'stare' < *kanakana* 'stare wildly'

*whākao* = *whakakao*<sup>17</sup> 'collect' < *kao* 'assembled'  
*whākapakapa* 'quivering of the hands' cf. *whakakakapa* 'brandish, flourish'  
*whākoekoe* = *whakakoekoe* 'tickle'  
*fakorekore* 'to deny' (Harlow 1987)  
*whākorekore* = *whakakorekore* 'cause not to be'  
*faakumu* 'to shut the mouth' (Harlow 1987)  
*whākuru* 'pelt' < *kuru* 'pelt' but *whakakuru* < *kuru* 'weary'  
*whāngongo* 'administer liquid food to someone' < *ngongo* 'suck' (but  
*whakangongo* 'neglect, pay no attention' < *ngongo* 'low born')  
*whāmutu* = *whakamutu* 'leave off'  
*whāngote* = *whakangote* 'suckle' (latter not in Williams 1971)  
*whāomoomo* = *whakaomoomo* 'tend a child'  
*whāpiko* 'make a snare or noose', *whakapiko* 'bend'  
*whāranu* = *whakaranu* 'mix (vtr)'  
*whārite* = *whakarite* 'make like, compare'  
*whātare(tare)* 'stoop over', *whakatare(tare)* 'lean forward'  
*whātika* = *whakatika* 'straighten'  
*whātoro* = *whakatoro* 'touch'  
*whātuturi* = *whakatuturi* 'be obstinate' < *turi* 'deaf'  
*whāunu* 'give to drink'

## APPENDIX 2

Passives of reduplicated stems:

*huti, huhuti* ~ *hūtia* 'hoist, haul up'  
*kakahu* ~ *kāhua* 'bite' South Island Maori (cf. Harlow 1987)  
*kukume, kume* ~ *kūmea* 'pull, drag'  
*kukuti, kuti* ~ *kūtia* 'draw tightly together, contract, pinch'  
*nonoke* 'struggle' ~ *nōkea* 'hustle him'  
*nonoti, noti* ~ *nōtia* 'pinch, contract'  
*papaki* ~ *pākia* 'to slap'  
*poki, popoki, pokipoki* ~ *pōkia* 'cover over'  
*pupuhi* ~ *pūhia* 'to blow or shoot'  
*tatari* ~ *tāria* 'wait' (*tāringa*, nom.)  
*tungi, tutungi* ~ *tūngia* 'kindle, set light to'  
*wau, wawau* 'foolish' ~ *wāuia* 'be discussed'

## APPENDIX 3

Some examples of the patterns of reduplication of trimoraic stems:

Simplex	Pattern 1	Pattern 2	Pattern 3	Pattern 4
* <i>arai</i>	<i>ārai</i> screen		<i>araarai</i> frequentative	

<sup>17</sup>Not in Williams (1971), but does occur in modern Maori.

<i>*hapai</i>	<i>hāpai</i> lift up		<i>hapahapai</i> frequentative	
<i>hawai</i> an edible fungus			<i>hawahawai</i> hillock in which weeds are buried	
<i>hoia</i> annoyed	<i>hohoia</i> annoyance			
<i>*hokai</i>	<i>hōkai</i> extended		<i>hokahokai</i> extend	<i>hōkaikai</i> extend
<i>horoi</i> wash			<i>horohoroi</i> <sup>18</sup> to wash, shower (vi)	
<i>kanono</i> <i>Coprosma australis</i>		<i>kānonono</i> <i>Coprosma australis</i>		
<i>kopū</i> blistered	<i>kokopū</i> type of gravel			<i>kōpūpū</i> having blistered skin
<i>?manea</i> talisman			<i>manemanea</i> satisfied	
<i>manei</i> reach out to		<i>mānenei</i> reach out to		
<i>mangungu</i> broken, chipped				<i>māngūngungu</i> gritty
<i>maoa</i> cooked, ripe	<i>mamaoa</i> steam		<i>maomaoa</i> first fruits of kumara	
<i>marara</i> scattered				<i>mārārara</i> rather scattered
<i>matiti</i> split, crack		<i>mātītiti</i> cracked, split		
<i>?matoru</i> benumbed	<i>mātoru</i> crowd	<i>mātoruru</i> thick, dull		<i>mātorutoru</i> benumbed
<i>merau</i>			<i>meramerau</i> <sup>19</sup>	
<i>mohua</i> yellowhead, k.o. bird	<i>momohua</i> yellowhead, k.o. bird			<i>mōhuahua</i> yellowhead, k.o. bird
<i>*mokai</i>	<i>mōkai</i> pet bird or animal		<i>mokamokai</i> pet bird or animal	<i>mōkaikai</i> pet bird or animal

<sup>18</sup>Not in Williams (1971), but cf. Moorfield (1988:145).

<sup>19</sup>No gloss in Williams (1971) for these two forms.

<i>*monehu</i>	<i>mōnehu</i> die, expire	<i>mōnenehu</i> indistinct	<i>mōnehunehu</i> indistinct	
<i>*moutu</i>	<i>momoutu</i> rifleman			<i>mōutuutu</i> rifleman
<i>pahū</i> explode	<i>papahū</i> resound, burst into flame			<i>pāhūhū</i> pop, crackle
<i>pakaru</i> broken	<i>pākaru</i> break in pieces (vtr)			<i>pākarukaru</i> break in pieces (vtr)
<i>pakē</i> creak			<i>pakepakē</i> creak	<i>pākēkē</i> creak
<i>pakeke</i> adult	<i>pākeke</i> adult			
<i>pakini</i> nip, pinch		<i>pākikini</i> pain, smart		<i>pākinikini</i> pain, smart
<i>panoko</i> k.o. fish	<i>papanoko</i> k.o. fish	<i>pānonoko</i> k.o. fish		<i>pānononoko</i> k.o. fish
<i>parau</i> false, baffled			<i>paraparau</i> baffled	
<i>paroro</i> threatening clouds	<i>pāroro</i> threatening clouds			
<i>parure</i> confused, abashed	<i>pārure</i> overcome, maltreat			
<i>*patai</i>	<i>pātai</i> ask		<i>patapatai</i> frequentative	<i>pātai tai</i> frequentative
<i>patō</i> knock			<i>patopatō</i> <sup>20</sup> type	<i>pātōtō</i> knock frequentative
<i>peau</i> to be turned away			<i>peapeau</i> turn aside (vi)	
<i>pirau</i> rotten	<i>pipirau</i> decayed		<i>pirapirau</i> decayed	
<i>*pokai</i>	<i>pōkai</i> roll up		<i>pokapokai</i> roll up	
<i>*potae</i>	<i>pōtae</i> hat		<i>potapotae</i> surrounding	

<sup>20</sup>Not in Williams (1971), but does occur in modern Maori.

<i>powhiwhi</i> tangled	<i>pōwhiwhi</i> several climbing plants		<i>pōwhīwhiwhi,</i> <i>pōwhiwhiwhiwhi</i> tangled
<i>ramene</i> assemble		<i>rāmemene</i> assemble	
<i>riaki</i> raise			<i>riariaki</i> raise
<i>riroi</i> twisted			<i>riroriroi</i> entangled
<i>riua</i> borne away			<i>riuriua</i> borne away
<i>tahō</i> yielding, weak			<i>tāhōhō</i> soft, pulpy
<i>takai</i> wrap up	<i>tākai</i> bandage (noun)		<i>takataakai</i> wind round
<i>tapahi</i> cut			<i>tapatapahi</i> cut in pieces
<i>*tarai</i>	<i>tārai</i> fashion, dress		<i>taratarai</i> fashion, dress
<i>tāu</i> turn away			<i>tātāu</i> turn about, vacillate
<i>tawē</i> noise			<i>tawetawē</i> noisy
<i>taweke</i> linger		<i>tāweweke</i> slow, dilatory	
<i>*tīhoi</i>	<i>tīhoi</i> wander, vagrant		<i>tīhotīhoi</i> wander aimlessly
<i>tinei</i> quench, extinguish			<i>tīhoihoi</i> wander aimlessly
<i>tunou</i> nod the head			<i>tinetinei</i> quench, extinguish
<i>tupehu</i> blustering			<i>tunotunou</i> bow repeatedly
<i>turaki</i> throw down			<i>whaka-tūpehupehu</i> bluster
<i>turua</i> beautiful			<i>turaturaki</i> throw down a number of things
			<i>turuturua</i> beautiful



<i>*whangai</i>	<i>whāngai</i> feed	<i>whangawhangai</i> charm	<i>whāngaingai</i> food to send visitors on their way
<i>wheau</i> be long (in time)		<i>wheawheau</i> a form of ritual	<i>whakawhēauau</i> delay (vi)

## BIBLIOGRAPHY

- ARMS, D.G., 1989, Problems in Fijian orthography. In Ray Harlow and Robin Hooper, eds *VICALI: Oceanic Languages: papers from the Fifth International Conference on Austronesian Linguistics*, 1-32. Auckland: Linguistic Society of New Zealand.
- BAUER, Winifred, 1981, Hae.re vs. ha.e.re: a note. *Te Reo* 24:31-36.
- ELBERT, Samuel H. and Mary Kawena PUKUI, 1979, *Hawaiian grammar*. University of Hawaii Press.
- HARLOW, Ray, 1987, *A word-list of South Island Maori*. *Te Reo* Monograph. Auckland: Linguistic Society of New Zealand.
- HAYES, Bruce and Maya ABAD, 1989, Reduplication and syllabification in Ilokano. *Lingua* 77:331-374.
- KRUPA, Viktor, 1966, *Morpheme and word in Maori*. The Hague: Mouton.
- MARANTZ, Alec, 1982, Re reduplication. *Linguistic Inquiry* 13:435-482.
- MOORFIELD, John C., 1988, *Te Kākano*. Auckland: Longman Paul.
- WILLIAMS, H.W., 1971, *A dictionary of the Maori language*. 7th edn. Wellington: Government Printer.

## A REANALYSIS OF PROTO POLYNESIAN NOUN PHRASE MARKING

S.P. HARRISON

### 1. INTRODUCTION: WHY RE-OPEN AN OLD ISSUE?<sup>1</sup>

#### 1.1 THE ERGATIVE-ACCUSATIVE DEBATE

The history of Polynesian nominal marking, particularly case marking, was the subject of considerable debate throughout the 1970s, beginning with Hohepa (1969). In that paper, Hohepa, following a suggestion by Ken Hale, proposed that Proto Polynesian (PPN) had an accusative case-marking typology, much like that of the contemporary Eastern Polynesian languages.<sup>2</sup> In that account, PPN transitive verbs appeared in either active or passive frames, as in contemporary Maori:

MAORI      *Ka inu te tangata i/ki te wai.*  
              TNS V ART man ACC ART water  
              The man drank the water.

*Ka inumia te wai e te tangata.*  
              TNS V+PASS ART water AG ART man  
              The water was drunk by the man.

The ergative pattern:

SAMOAN    *Saa inu(mia) e le tagata le vai.*  
              TNS V(+CIA) ERG ART man ART water  
              The man drank the water.

in Tongic and in the Samoic-Outlier languages is held to be the result of a drift-like generalisation and, ultimately, a literal unmarking of the passive in those languages.

---

<sup>1</sup>I am pleased to dedicate this work to my teacher and friend, Prof. G.W. Grace. I have learned many things from him, not the least of which is how one should endeavour to cope with the pressures and frustrations of academic life. An early version of this paper was presented to the Honolulu Austronesian Circle in late 1983, after which it lay fallow for many years (not long enough, some might say!). I should like to thank Anthony Aristar for insightful discussion of the relationship between noun semantics and case marking, and Bob Blust for valuable critical comment on an earlier draft.

<sup>2</sup>Abbreviations used are as follows: ABS – absolute, ACC – accusative, AG – agent, ART – article, BENF – benefactive, CN – common noun, DIR – directional, ERG – ergative, INST – instrument, LOC – locative, NSP – non-specific, PASS – passive, PERF – perfective, PN – proper noun, POSS – possessive, PRED – predicative, PST – past, SRC – source, TNS – transitive.

Clark (1976:68ff.) objects to that reconstruction on the grounds that (1) it necessitates a change cutting across established Polynesian subgroups, that (2) Hohepa's account of the progress of the change, in terms of 'drift', is unmotivated, and that (3) it fails to account for all the facts of contemporary Polynesian case marking.

Clark's alternative account is initially appealing. He recognises a contrast in Polynesian between what he terms 'A- and B-verbs'. The former have experiencer subjects or locative objects, and are perhaps conceptually less transitive than the latter, which are 'typical' transitive verbs involving an agent and a patient. Clark then reconstructs three case-marking patterns for PPN:

Pattern 1:	V	S	<i>i/ki</i>	O
Pattern 2:	V+Suff	e	S	O
Pattern 3:	V	e	S	O

Patterns 1 and 2 were available for A-verbs, and patterns 2 and 3 for B-verbs. Pattern 1 is a transitive reinterpretation of an Oceanic intransitive construction (where *\*i* and *\*ki* are oblique prepositions). 'Suff' is the pair PPN *\*-Ci* and PPN *\*-a*, the former reflecting the Proto Oceanic transitive and the latter, the Proto Oceanic third person singular object suffix. The locus of internal Polynesian change, in Clark's analysis, is Eastern Polynesian. In those languages, pattern 3 was for the most part lost, and pattern 1 was extended from A-verbs to all transitive verbs (though in a number of languages, he claims, it remains more common with A-verbs).

The issue was not laid to rest by Clark's study, however. Chung (1978) presented a sophisticated formal argument for Hohepa's original claim that PPN was accusative. From the perspective of an early version of relational grammar, she argues that Clark's pattern 1 was definitely transitive (rather than oblique) in PPN, even for what she calls middle verbs (Clark's A-verbs). But her arguments have less to do with evidence that NPs flagged by *i* in Polynesian are direct objects than with evidence that subjects in clauses with *i* NP undergo rules otherwise restricted to *transitive* subjects. Therefore, the clauses in which such subjects appear are *transitive* clauses, *i* NPs are direct objects, and *i* is an accusative marker. It is not clear to me, given subsequent revisions to the formal model, how her arguments have stood the test of time.<sup>3</sup>

## 1.2 THE PPN NOUN PHRASE

How can one summarise the debate surrounding PPN case marking in the 1970s? First, I think, by observing that the debate seemed to have little to do with *nominal case marking* at all. With the exception of Chung's formal argument regarding the status of PPN *\*i*, the focus of the debate was the history of transitivity-related verbal suffixes in Polynesian, and not NP internal case marking at all. It was taken for granted that the basic structure of the PPN noun phrase was:

Prep Article Noun

that the prepositions marked case, and that the articles marked definiteness or number.

In the present paper, I will attempt to redress the balance somewhat by re-examining Polynesian NP marking itself. As my point of departure, I take Clark's (1976) reconstruction of Proto Polynesian prepositions and articles. I provide a critical evaluation of those reconstructions in section

<sup>3</sup> Chung was at least as concerned, if not more concerned, with arguments that PPN *\*-Cia* was a passive suffix as with arguments that *\*i* was an accusative marker. I cannot deal with those arguments here, however.

2. Section 3 treats noun marking in Tongic in some detail. Finally, in section 4, I make some new proposals regarding PPN nominal case marking and consider some lingering problems.

2. PPN PREPOSITIONS AND ARTICLES

2.1 THE INVENTORY

Clark (1976:36-61) reconstructs the following prepositions and articles for Proto Polynesian:

PREPOSITIONS		ARTICLES	
*∅	absolutive	*te	definite
*e	ergative	*sa	indefinite
*i	locative, cause	*a	pronominal/proper
*ki	directional	*∅	plural
*ko	topic		
*mai	source		
*a/*o	possessive		

PPN *\*ki* also marks instruments and recipients. As noted above, both PPN *\*i* and *\*ki* appear as ‘accusative’ markers for Clark’s A-verbs. Clark (pp.47f.) notes that apparent reflexes of *\*ko* occur in a variety of functions, but argues that most of these can be derived from an earlier topic or predicate nominal flagging function. He also reconstructs PPN *\*ʔaa* ‘paucal’ following either the PPN definite or the plural article.

2.2 CLARK’S ACCOUNT OF PPN ARTICLES

2.2.1 PPN *\*te* AND *\*sa*

Clark’s account of post-PPN developments in the form of his reconstructed article system is marred, in my view, by its frequent appeal to irregular sound changes, as the following diagrams suggest (where PTO represents Proto Tongic and PNP represents Proto Nuclear Polynesian):

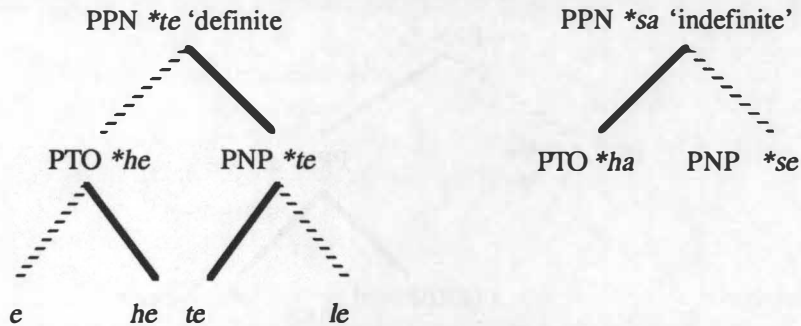


FIGURE 1: PPN DEFINITE AND INDEFINITE ARTICLES

In Figure 1, broken lines indicate irregular phonological developments:

- (a) PPN *\*t* is regularly [s] in Tongic before reflexes of PPN *\*e*. Tongan and Niuean *he* suggest a PPN *\*se*, rather than *\*te*.

- (b) In order to account for the distribution of the Tongan and Niuean definite articles *e* and *he* after various synchronic prepositions (see section 3), Clark assumes that the initial consonant of the definite article, whatever it was, was irregularly lost *after a back vowel* in Proto Tongic.
- (c) PPN *\*t* is reflected as /l/ in the definite article *le* of Samoan and East Futunan. While this is not an isolated reflex (being attested also in reflexes of PPN negative morphemes – see Clark 1976:85ff.), it is not a regular one.
- (d) PPN *\*a* is irregularly reflected as PNP *\*e* in Nuclear Polynesian reflexes of the indefinite article. (Clark reconstructs PPN *\*sa* rather than *\*se*, on account of the Proto Eastern Oceanic reconstruction *\*sa* ‘one’.)

### 2.2.2 PPN *\*a*

Clark (1976:58ff.) reconstructs a personal/pronominal article PPN *\*a* on the basis of the following reflexes:

- (a) proper and pronominal forms of the prepositions *\*i*, *\*ki* and *\*mai*, suggestive of *\*i a*, *\*ki a* and *\*mai a*, throughout Polynesia;
- (b) personal and pronominal (nominative) articles in a number of Nuclear Polynesian languages;
- (c) the Tongic absolutive marker PTO *\*qa*;
- (d) plural articles in several Samoic languages.

The first two of these functions he assumes to be derivative of the PPN form; the last two, he claims, are idiosyncratic post-PPN developments. In pre-PPN, Clark conjectures *\*a* was a simple personal/pronominal article. By PPN it was restricted to NPs with the prepositions *\*i*, *\*ki* or *\*mai*, or with no preposition. In order to account for this distribution, he postulates yet another irregular phonological change, a morphophonemic rule deleting *\*a* after prepositions ending in a non-high vowel (that is, after the possessive markers *\*a/\*o*, *\*ko* and *\*e*). The remaining reflexes of *\*a* fused with a preceding preposition. These developments can be diagrammed:

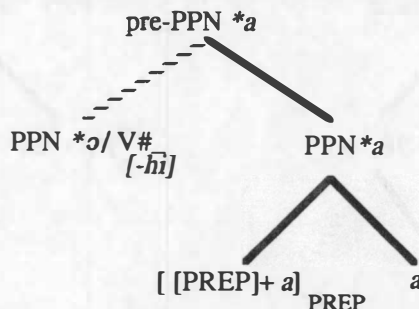


FIGURE 2: PPN *\*a* ‘PERSONAL, PRONOMINAL’

In a few Nuclear Polynesian languages (Maori, for example), PPN *\*a* survives as a proper noun nominative (and fossilised on some personal pronouns). In most of the remaining better-known Nuclear Polynesian languages (like Samoan, Tahitian and Hawaiian), ‘free’ *\*a* has been lost, but in a few Samoic languages, it (under Clark’s analysis) has been reanalysed as a plural article. He

conjectures that, following the fusion of *\*a* with a preceding preposition, 'free' *\*a* was reinterpreted as a sort of default noun marker, replacing PPN *\*∅* as a plural article. He does not account for its subsequent loss in its original environment, with nominative proper names. Clark assumes a similar development in Tongic, in which the 'free' reflex of PPN *\*a* was reinterpreted as a nominative marker for both proper and common nouns. (A similar development appears in some Samoic languages, Rennellese for example – see Clark 1976:61.)

### 3. NOMINAL MARKING IN TONGIC

#### 3.1 NIUEAN PREPOSITIONS AND ARTICLES

The Niuean nominal-marking system is described in relative detail in Seiter (1980:27-60). The following table, adapted from Seiter (p.37), summarises noun marking in Niuean:

TABLE 1: NOMINAL MARKING IN NIUEAN

	ABS	ERG	LOC	DIR	INST	COM	BENF	POSS	PRED	SRC
CN	<i>e</i>	<i>he</i>	<i>he</i>	<i>ke he</i>	<i>akie</i>	<i>moe</i>	<i>mae</i>	<i>he</i>	<i>ko e</i>	<i>mai he</i>
PN	<i>a</i>	<i>e</i>	<i>i(a)<sup>4</sup></i>	<i>ki(a)</i>	<i>akia</i>	<i>mo</i>	<i>ma</i>	<i>a/ha</i>	<i>ko</i>	<i>mai i</i>

What is immediately striking about Niuean is that:

- (a) case is always marked

that is, there is no zero case marking, and that:

- (b) there is only one prenominal slot for absolutive, ergative and locative nouns

that is, preposition and article are not distinguished for these three cases.

There are, however, some exceptions to each of these claims. Absolutive *e* is optional before a preposed possessive, cardinal number (or the quantifiers *loga* 'many' or *gahoa* 'few'), and *ha* 'non-specific' or *taha* 'indefinite':

NIUEAN    *Nofo (e) haana a        tupuna        fifine i        Avatele.*  
 live    ABS his        POSS grandparent female LOC Avatele  
 His grandmother lives in Avatele.

*Mate tuai (e) ua e kulī.*  
 die    PERF ABS two dog  
 Two dogs died.

*Ai fia loto (e) ha tagata i        Niu Silani        ki a        ia.*  
 not want want ABS NSP man    LOC New Zealand to ART him  
 Nobody in New Zealand wanted him.

<sup>4</sup>Blust (pers. comm.) reminds me that at least three distinct (non-affixal) morphemes of the form *\*i* have been reconstructed (at various levels in the Austronesian tree, at least as far down as POC): a locative preposition, a personal article and a genitive marker. He suggests that PPN *\*i* is the result of a (partial) merger of these morphemes. That may well be, though I've never been particularly disposed toward projecting grammatical homophones indefinitely far back in time. I would prefer attempting to motivate a single reconstruction, from which later homophones can be derived. But this issue can safely be sidestepped in the current context.



Locative personal names, as distinct from place names, take *a* when flagged by the prepositions *i* or *ki*:

*Malona tuai e kapiniu ē i a Maka.*  
 broken PERF ABS dish this LOC ART Maka  
 This dish got broken on account of Maka.

*Nofo e taokete haana i Hakupu.*  
 live ABS brother his LOC Hakupu  
 His brother lives in Hakupu.

*Ne fakafano e tohi ki a Sione.*  
 PST send ABS letter to ART Sione  
 The letter was sent to Sione.

*Ne fakafano e tohi ki Niuē.*  
 PST send ABS letter to Niue  
 The letter was sent to Niue.

Note also that instances of the non-specific article *ha* provide a class of exceptions to generalisation (b) above. The non-specific *ha* is one of a small number of items that may intervene, in Niuean, between a case marker and its noun. Others include *tau* 'plural', *kau* 'group (of people)', *lafu* 'group (of kin)' and *nā* 'pair'. The indefinite article *taha* 'some (individual)' is similar to *ha* in that the absolutive marker *e* is optional before it.

Pronominal and proper noun marking is identical in Niuean, except in the possessive and the (related) benefactive, but these details are irrelevant for the purposes of this discussion (see Seiter 1980:34-36).

### 3.2 TONGAN PREPOSITIONS AND ARTICLES

Proper/pronominal absolutive, ergative and locative markers all have an initial glottal stop in Tongan, and the locative and directional prepositions have pre-pronominal forms *'iate* and *kiate*. Otherwise, in most relevant respects, Tongan proper/pronominal noun marking is identical to Niuean:

TABLE 2: PROPER/PRONOMINAL NOUN MARKING IN NIUEAN AND TONGAN

	ABS	ERG	LOC	DIR	PRED	SRC
NIUEAN	<i>a</i>	<i>e</i>	<i>i (a)</i>	<i>ki (a)</i>	<i>ko</i>	<i>mai i</i>
TONGAN	<i>'a</i>	<i>'e</i>	<i>'i (a)</i>	<i>ki (a)</i>	<i>ko</i>	<i>mei (a)</i>

Note that the absolutive *'a* is rare with pronouns and demonstratives in Tongan.

Common noun marking differs markedly, however:

TABLE 3: COMMON NOUN MARKING IN NIUEAN AND TONGAN

	ABS	ERG	LOC	DIR	
NIUEAN	<i>e</i>	<i>he</i>	<i>he</i>	<i>ke he</i>	(SPECIFIC)
TONGAN	<i>(a) e</i>	<i>'e he</i>	<i>(i) he</i>	<i>ki he</i>	
NIUEAN	<i>(e) ha</i>	<i>he ha</i>	<i>he ha</i>	<i>ke he ha</i>	(NON-SPECIFIC)
TONGAN	<i>ha</i>	<i>'e ha</i>	<i>(i) ha</i>	<i>ki ha</i>	

Unlike Niuean (but like other Polynesian languages), Tongan has two distinct prenominal slots in common noun phrases. The first is occupied by a prepositional case marker, and the second by an article. The Tongan specific article is *e* in the absolutive, and *he* otherwise; the non-specific, like Niuean, is *ha*. The Tongan absolutive case marker '*a*' is optional with specific common nouns, and disallowed with non-specific; the locative '*i*' is always optional in common noun phrases. For other details of Tongan nominal marking, see Churchward (1953).

### 3.3 CLARK'S ACCOUNT OF NIUEAN NOUN MARKING

As noted above, Clark's (1976) account of the development of nominal marking in Tongic rests on two hypotheses. The first is the irregular development of PPN *\*te* 'definite article' into pre-PTO *\*se* > PTO *\*he/\*e*, the former after front vowels (*\*i* or *\*e*) and the latter after back vowels (*\*a* and *\*o* being the only relevant cases). Clark is not explicit, as Seiter (1980:342 fn.4) is, that the putative Tongic reflexes are *specific* articles, rather than *definite*. Clark's second hypothesis is that the PPN personal/pronominal article *\*a* was reinterpreted as an absolutive marker *\*a* in Proto Tongic,<sup>5</sup> and in that way spread from proper to common nouns.

Subsequent developments in Tongan are straightforward. Either the development of *\*a* as an absolutive marker is still in progress in Tongan, accounting for its optionality in specific common noun phrases and its absence in non-specific contexts, or it has begun to decline in those contexts. The former is perhaps more plausible on the basis of Tongan evidence alone. Clark, however, offers no account of the optionality of the Tongan locative '*i*' in common noun phrases or before demonstratives (see Churchward 1953:116).

The derivation of Niuean common noun phrase marking from this Proto Tongic system requires more ingenuity, however. Clark hypothesises yet another irregular phonological change to account for synchronic specific common noun phrase marking in Niuean – the loss of #V# prepositions before reflexes of pre-Tongic *\*se*.<sup>6</sup> The historical development of Niuean nominal marking under Clark's account can then be described in terms of the following sequence:

TABLE 4: DEVELOPMENT OF NIUEAN PRO-/PROPER NOUN PHRASE MARKING

	ABS	ERG	LOC	DIR
pre-PPN	<i>*∅ a</i>	<i>*e a</i>	<i>*i a</i>	<i>*ki a</i>
PPN	<i>*∅ a</i>	<i>*e ∅</i>	<i>*i a</i>	<i>*ki a</i>
PTO	<i>*a ∅</i>	<i>*e</i>	<i>*i a</i>	<i>*ki a</i>
NIU	<i>a</i>	<i>e</i>	<i>i a</i>	<i>ki a</i>

The evolution of Niuean pronominal/proper noun marking offers no surprises in Clark's account. One need only postulate the loss of the proper article *\*a* after non-high vowels and the (in this case vacuous) reinterpretation of 'free' *\*a* as an absolutive marker.

<sup>5</sup>The Tongan reflexes of PPN *\*e* and *\*i* have a glottal stop, as do at least some of the reflexes of PPN *\*a*. Since other Polynesian languages with non-zero reflexes of PPN *\*ʔ* do not show glottal stop in these forms, it is likely that their Tongan form is an innovation. It is not clear whether that innovation should be attributed to Proto Tongic, since Niuean reflects PPN *\*ʔ* as zero. Nor is it clear under what conditions the Tongan forms acquired their glottal stop, since no glottal stop appears in post-prepositional *a* before proper nouns and pronouns.

<sup>6</sup>Loss of single-vowel prepositions must be restricted to this environment, since reflexes of these prepositions remain before reflexes of PPN *\*a*.

TABLE 5: DEVELOPMENT OF NIUEAN SPECIFIC COMMON NOUN PHRASE MARKING

	ABS	ERG	LOC	DIR
PPN	* $\emptyset$ te	*e te	*i te	*ki te
pre-PTO <sub>1</sub>	* $\emptyset$ se	*e se	*i se	*ki se
pre-PTO <sub>2</sub>	*a he	*e he	*i he	*ki he
PTO	*a e	*e he	*i he	*ki he
NIU	e	he	he	ke he

The evolution of specific common noun phrase marking is not quite as straightforward. First, one requires the irregular development of PPN \*te into \*se in pre-Tongic. In order to obtain the PTO absolutive specific article \*e, one must assume the following chronology:

- PPN \*a 'personal/pronominal article' > PTO \*a 'absolutive';
- pre-PTO \*se 'specific article' > PTO \*e before back vowels;
- PTO single segment prepositions >  $\emptyset$  before articles.

If the personal article had not been generalised to absolutive function ((a) above) early, then the morphophonemic solution to the e/he alternation ((b) above), as it is usually described, fails. Since the e alternate is assumed to have arisen when preceded by a non-high vowel final preposition, some appropriate preposition must have been present (unless one assumes that PTO \*he remains unchanged only after high vowel final prepositions, and not elsewhere).

Finally, note that the history provided by Clark gives no account whatsoever of Niuean non-specific common noun marking. Unless one assumes a period in which Niuean non-specific noun phrases, under Clark's analysis, had *two* articles (reflecting a putative sequence PPN \*te sa), one can only claim that Niuean specific noun phrase marking has spread relatively recently to non-specific noun phrases, replacing whatever role marking the latter had earlier carried. Such an account becomes strained, I think, if one is to claim that, synchronically, Niuean he and e are best viewed as articles rather than case markers.

#### 4. POLYNESIAN NOMINAL MARKING: AN ALTERNATIVE HISTORY

##### 4.1 THE EVALUATION OF LINGUISTIC HISTORIES

Let me begin by isolating four criteria in terms of which alternative linguistic histories can be evaluated. I will term these (a) simplicity, (b) completeness, (c) historicity and (d) plausibility. The *simplicity* and *completeness* criteria are straightforward. The first is simply Occam's Razor. The second says that the more state descriptions that are related in a history, the better that history.

*Historicity*, simply put, is the conviction that everything comes from somewhere. As a criterion for evaluating alternative histories of some item, it ranks those histories in direct relation to how far back in time or how high up the tree they trace the item in question. In short, the older, the better! The historicity criterion encourages the historian to favour histories that do *not* involve 'spontaneous coinage' over those that do. (The nature and justification of the historicity criterion is considered in greater detail, with respect to a particular instance of its use, in footnote 13 below.)

The criterion labelled *plausibility* might have been termed *naturalness* or *explanatory value*. There are two ways in which the criterion of plausibility can be brought to bear on the evaluation of linguistic histories. The first is the demand that any single change (relation between components of

adjacent historical states) conform to some (theoretically-sanctioned) view of what a possible change is. Plausibility in this sense is perhaps the most nebulous of the evaluation criteria, given that there is no complete, accepted, and explicit theory of linguistic change. Using plausibility to evaluate changes is to a great extent a matter of professional experience.

The second application of a plausibility criterion is in the evaluation of reconstructions – state descriptions in a history that are not attested/observable but are inferred from the evidence of subsequent states in the same historical system. Independent of any other criteria by which histories (whether involving reconstruction or not) are evaluated, most historical linguists would claim that a reconstructed state description must conform to the theoretical constraints on (synchronic) linguistic state descriptions; that is, a reconstruction must look like (a description of)<sup>7</sup> natural language.

Considerations of both simplicity and plausibility typically lead to comparative reconstructions that resemble one language or subgroup more than others. (The application of Teeter's Law is thus not without metatheoretical justification.) A reconstructed system that resembles an attested system is bound to be plausible, and the subsequent history of that system is bound to be simpler than possible alternatives, since at least one subsequent step is largely an identity map. This is certainly true of Clark's account of Polynesian nominal marking, to the extent that it is a generalisation on attested Nuclear Polynesian nominal marking.

#### 4.2 AN EVALUATION OF CLARK'S 'STANDARD' ACCOUNT

The PPN nominal-marking system Clark reconstructs is grounded in synchronic Nuclear Polynesian nominal-marking systems. Its salient characteristics are:

- (a) two distinct prenominal marker slots: *preposition* and *article*;
- (b) a distinction between *definite* and *indefinite* articles;
- (c) a contrast between *common* and *proper/pronominal* marking, reflected in the choice of article.

PNP nominal marking is identical to PPN, except for the form of the indefinite article (PPN *\*sa*, PNP *\*se*) and the fusion of PPN *\*a* 'personal article' with a preceding preposition in PNP. Post-PNP developments include the development of irregular reflexes of the PPN/PNP definite article *\*te* in some Samoic languages and the various changes affecting reflexes of PPN *\*a* 'personal/pronominal article', outlined above. The major change, in Clark's proposal, was the change toward accusative case-marking typology in Eastern Polynesian and in some Samoic-Outlier languages. The opposing camp, for whom PPN was accusative, does not dispute Clark's reconstructions of the internal structure of the PPN noun phrase, nor of the form of its prepositions and articles. It is only with respect to the functions of two of the reconstructed prepositional forms that there is any disagreement between the two camps.

The standard account does not fare well, particularly with respect to the criteria of simplicity and of plausibility, in deriving the synchronic Tongic nominal-marking systems. The standard account requires three lexical-specific changes in form (PPN *\*te* (pre-PTO *\*se*) > PTO *\*he*, PTO *\*he* > PTO *\*e* after a back vowel, and the loss of #V# prepositions before the specific article in Niuean), in

<sup>7</sup>It is not always easy to distinguish between a reconstruction that looks like natural language and one that conforms to some current theoretical descriptive standard. Some practitioners might, for example, regard a reconstruction more highly if it can be represented in terms of government-and-binding theory, or metrical phonology, or whatever.

addition to the loss of pre-PPN \*a 'personal/pronominal article' after non-high vowel final prepositions that Clark postulates to account for the distribution of the personal article in his PPN system. Not only are these changes less than completely general, but, as far as my own intuitions are concerned, most of them also lack plausibility. Clark's history of the development of Tongic nominal marking is simply a description of how one might arrive at Tongan and Niuean nominal marking from a PPN reconstruction grounded in Nuclear Polynesian. The Tongic data themselves do not appear to have had much influence on the PPN reconstructions.

#### 4.3 PREMISES FOR A NEW RECONSTRUCTION

The history of Polynesian nominal marking offered here is based on two premises:

- (a) As far as the shape of reconstructed lexical items is concerned, what you see should be what you get. Nonce form changes of the sort employed in the standard account are to receive strong negative weight in evaluating alternative histories. With respect to the forms in question, this premise casts doubt on the cognacy of:
- i. PNP \**te* and PTO \**he*
  - ii. PNP \**se* and PTO \**sa*

and on any allomorphic relation between PTO \**e* and PTO \**he*, as claimed in the standard account.

Rather, it suggests that one explore the possible cognacy of:

- i. PNP \**e* 'ergative' and PTO \**e*
  - ii. PNP \**se* 'indefinite article' and PTO \**he*
- (b) Tongic might provide as good a basis, if not a better basis, for the reconstruction of PPN nominal marking as does Nuclear Polynesian. No matter how many languages each subgroup contains, there is no reason, a priori, to give more weight to one than to the other in reconstructing PPN.

#### 4.4 A REANALYSIS OF NIUEAN NOMINAL MARKING

##### 4.4.1 NON-OBLIQUE NOUN PHRASES

The history to be developed here rests heavily on the synchronic analysis of Tongic, particularly Niuean, nominal marking. Consider first the marking of non-oblique noun phrases in Niuean. Such noun phrases are marked in one of three ways:

- (a) by *a*, if the noun is proper/pronominal and absolutive;
- (b) by *he*, if the noun is common and ergative;
- (c) by *e* otherwise.

For the non-oblique cases, the present account of Niuean nominal marking assumes a single prenominal slot filled by a marker that is perhaps more article than preposition, but is actually a bit of both (much like the construction markers of Philippine languages). The choice of marker is governed by two properties, the head noun's *class* and its *case* (grammatical relation). The primary case contrast is between *absolutive* and *ergative*. For reasons that will become apparent, it is also convenient to consider *locative* to be a (third) non-oblique case contrast. The primary class division



is between *proper* and *common* nouns. The former can be divided into three subclasses: *personal*, *pronominal* and *locative*. Niuean, like other Polynesian languages, has a class of nouns that are *locative* in terms of their inherent semantics. As in other Polynesian languages, Niuean locative nouns include place names, nouns denoting directions, and orientationally defined parts (such as 'inside', 'top', 'back').<sup>8</sup>

Relativising to noun class, the distribution of non-oblique noun markers in Niuean can be described as follows:

- (a) *e* flags the *unmarked* case for both common nouns and proper non-locative nouns – for common nouns, the unmarked case is *absolutive*; for proper (non-locative), it is *ergative*;
- (b) *i* flags the *unmarked* (locative) case for locative nouns;<sup>9</sup>
- (c) *a* flags the *marked* (absolutive) case for proper nouns;
- (d) *he* flags the *marked* (ergative or locative) cases for common nouns.

This analysis assumes that each noun class has a particular unmarked grammatical relation (case), and that noun-marker choice is to some degree governed by this property. (There may also be a sense in which the absolutive case can be described as, in some absolute sense, unmarked, in that every clause has at least one absolutive nominal. This property of the absolutive case may have had some influence on the history of PPN \*a – see section 4.5.)

TABLE 6: NOUN CLASS AND NON-OBLIQUE NOMINAL MARKING IN NIUEAN

		NOUN CLASS					
		COMMON		PROPER		LOCATIVE	
CASE	unmarked	ABS	<i>e</i>	ERG	<i>e</i>	LOC	<i>i</i>
	marked	ERG/LOC <i>he</i>		ABS	<i>a</i>	LOC	<i>i a</i>

#### 4.4.2 OBLIQUE NOUN PHRASES AND TWO-SLOT NOMINAL MARKING

As observed earlier (section 3.1), non-oblique (absolutive, ergative and non-personal locative) specific noun phrases in Niuean have the structure M N, where N is a noun and M is proclitic noun marker selected on the basis of both case and inherent referential semantics. The 'proto-typical' Polynesian noun phrase, with prenominal slots for both a preposition and an article, is evidenced only in Niuean oblique noun phrases. From Table 1, one observes that, for common nouns, the directional marker *ke* and the source marker *mai* are followed by the 'article' *he*, while other Niuean oblique case markers (instrumental *aki*, comitative *mo*, benefactive *ma* and predicative *ko*) take *e*. Clark offers a historical phonological account of the distribution of Niuean *e/he* after oblique case markers, where the consonant of the article *he* was lost after prepositions ending in a back vowel<sup>10</sup>.

<sup>8</sup>Clark (1976:54ff.) terms these the L-class. See Harrison (1989) for a discussion of a similar class in Gilbertese and other Micronesian languages.

<sup>9</sup> Locative case *is* marked for pronouns and personal names. For such nouns, the proper locative marker *i* (and the proper directional *ki*) require a noun in the absolutive case.

<sup>10</sup> Incorrectly predicting \**aki he*, rather than the observed *aki e*.



Let me suggest another account. The oblique locative markers *ke* (*ki* with proper nouns) 'to' and *mai* 'from' require a noun that is marked *locative*, while the other oblique case markers take nouns that are marked as *absolutive*. On this account, one expects common nouns to be preceded by the common locative *he*, and proper nouns or pronouns to be preceded by the proper locative *i* after the directional and source oblique markers. These predictions are borne out except in the case of the proper/pronominal directional *ki*, which does not require a noun introduced by *i* (perhaps because the result would be two adjacent identical syllabic segments). Parallel predictions follow for the non-locative oblique markers, which one expects to be followed by either the absolutive *e* or *a*, depending on the nature of the noun. These last predictions are less accurate; only the instrumental *aki* requires a following proper noun or pronoun to be marked absolutive; other non-locative obliques precede proper nouns or pronouns directly.

Thus, while Niuean does have noun phrases that follow the canonical Polynesian pattern, with two slots preceding the noun itself, it is perhaps not accurate to label those slots 'preposition' and 'article', in the sense in which those categories are usually understood. The occupants of the Niuean 'article' slot are drawn from the same set as the Niuean non-oblique noun markers. Indeed, with the exception of those proper noun non-locatives noted above, one might perhaps best describe the structure of the Niuean two-slot noun phrase as a noun phrase NP preceded by an oblique role marker OBL,<sup>11</sup> where NP has the structure M N described above.

#### 4.4.3 NON-SPECIFIC NOUN PHRASES

The synchronic status of the Niuean non-specific article *ha* (see sections 3.1 and 3.2) deserves some comment here. Niuean *ha* regularly follows one of the common noun markers *e* or *he*. It is only to be regarded as an article in a preposition^article^noun construction if *e/he* are prepositions, an inappropriate label, as already suggested. In oblique noun phrases, *ha* follows both the oblique and the common noun marker:

NIUEAN      *Onoono e      tama ke he      ha gata.*  
                  look    ABS child to LOC a    snake  
                  The child is looking at some snake.

On the basis of the information available to me concerning its syntax and semantics, I would prefer to analyse Niuean *ha* as one of a distinct set of pronominal number markers, syntactically and semantically distinct from the common noun markers *e/he*.

#### 4.4.4 SUMMARY

In the analysis proposed here, the Niuean noun phrase does not have the preposition (case)^article^noun structure typical of Polynesian languages, and proposed for Niuean itself by Seiter (1980:27). Rather, I suggest that the structure:

(oblique-marker) noun-marker (number-marker)

is more revealing. Cognates of the prepositions of other Polynesian languages come from both the oblique-marker and noun-marker sets in Niuean. Cognates of the articles of other Polynesian

<sup>11</sup> It is not obvious that members of OBL should be regarded as prepositions, in the usual sense of that term. It is more likely that they are verbs or conjunctions that are in the process of developing into prepositions.

languages can likewise be found in the Niuean noun-marker set. But the Niuean noun markers are neither prepositions nor articles, but a distinct category with some of the properties of both.

#### 4.5 A NEW RECONSTRUCTION OF PPN NOUN MARKING

##### 4.5.1 THE RECONSTRUCTED SYSTEM

I assume the Proto Polynesian noun-marking system is more transparently reflected in contemporary Tongic (particularly Niuean) than in Nuclear Polynesian. As in Niuean, the terms preposition and article are not particularly appropriate labels for the class of prenominal noun markers I will reconstruct for PPN. I prefer the neutral term *noun marker*. PPN non-oblique noun phrases had a single obligatory prenominal slot, occupied by a noun marker selected in accordance with both the case and the class of the head noun. (Though it is simplest to regard the case-marking system as ergative, that label is somewhat misleading in view of the theory of case markedness outlined in section 4.4.1). The noun markers I reconstruct are:

- \*se marked common noun marker
- \*e unmarked (common and proper) noun marker
- \*a marked proper noun marker<sup>12</sup>
- \*i unmarked locative noun marker

In oblique noun phrases, an oblique marker preceded the noun marker. The choice of noun marker in oblique noun phrases depended on the role flagged by the oblique marker. The oblique markers reconstructed include:<sup>13</sup>

- \*ki goal
- \*mai source
- \*aki instrument
- \*ko topic

(I ignore possessive marking and other related oblique markers.) Assuming that Niuean is conservative, it is problematic that:

- (a) some oblique markers, including NIU *ko* 'topic', do not take a noun marker with proper nouns, and that;

<sup>12</sup> This analysis of PPN \*a, as a marked-case proper noun marker, may also be reflected in the use of \*a as a marker of 'dominant' possession.

<sup>13</sup> Unlike Clark, I reconstruct \*aki for PPN. That decision is motivated by the historicity criterion considered in section 4.1 and by the fact that the item is known to have a long history in Austronesian. I assume \*aki was lost in Nuclear Polynesian.

In comments on an earlier version of this paper, in which the *historicity criterion* was defined rather more formally in section 4.1 than it is in the present version (and perhaps more opaquely as well!), Blust (pers. comm.) objects that, in the present context, the historicity criterion is simply what he has called 'reconstruction from the top down'. Let me make clear that there is a *co-ordinate* conjunction in the second sentence of the preceding paragraph, and that the historicity criterion is indeed distinct from any 'top-down' reconstruction argument. If anything, the historicity criterion is a 'bottom-up' criterion, since it holds that morphemes (and particularly *grammatical* morphemes – if I can permit myself the use of that vague term), are *less* likely to have risen whole from the foam at some node in a family tree than they are to have a longer history. One might object that the historicity criterion licences the attribution of some morpheme to any or all nodes in a family tree above the lowest node for which it can be reconstructed on the basis of extant reflexes. And, in its unqualified form, indeed it might. It argues that, in general, loss/replacement of a (grammatical) morpheme is more likely than spontaneous coinage. In the case in question, I would not be in the least uncomfortable reconstructing PPN \*aki even if there were no evidence of the form outside Polynesia. Just how high up the tree one tries to project an item, using the historicity criterion, depends on at least three factors: the geometry of the tree, the shape of the reconstructed system, and one's 'theory' of grammatical/semantic change. In practice, I use the historicity criterion as a prod, to encourage myself to reconstruct *systems*, not just items, and to seek plausible histories by which items of diverse synchronic function might plausibly be linked.

- (b) the Niuean locative and directional oblique markers *i* and *ke* take (marked) absolutive pronouns and personal names.

I assume the latter to have been true of PPN as well, since that pattern is reflected in all Polynesian languages. One might want to assume that PPN common and proper nouns behaved differently after markers like PPN *\*ko* 'topic', as in contemporary Niuean, or that all PPN nouns required/did not require a noun marker after markers like *\*ko*. The decision depends on one's views regarding the benefits of symmetry in reconstructions.

In addition, I reconstruct a set of PPN noun number markers, including the (non-specific) number marker *\*sa* 'one, some'.

#### 4.5.2 CHANGES TO THE PPN NOUN-MARKING SYSTEM

##### 4.5.2.1 TONGAN

The principal post-PPN syntactic change in noun marking was the generalisation of a 'two-slot' pattern (reconstructed for PPN by Clark); that is, the emergence of distinct preposition and article categories and syntactic positions in all Polynesian languages but Niuean. I assume that PPN oblique noun phrase marking (and locative marking for non-locative proper nouns) served as a model for that pattern in PPN.

The Tongan system can be derived from the PPN reconstruction by, first, reanalysing (non-oblique) proper noun markers as prepositions and extending them to common nouns. In the resulting system, simple proper noun phrases are of the form:

prep PN

while simple common noun phrases are of the form:

(prep) ART CN

where the former common noun non-oblique markers are being reanalysed as (specific) articles. The emerging article class in Tongan is then augmented by the inclusion of (the Tongan reflex of) the PPN number marker *\*sa*.

##### 4.5.2.2 NUCLEAR POLYNESIAN

The noun-marking system of Nuclear Polynesian can be derived by extending the changes giving rise to the Tongan system. The two main Nuclear Polynesian innovations were:

- (a) modification/restriction of the function of the pre-PNP absolutive marker *\*a* (accompanied in Eastern Polynesian by the move toward accusative case marking);
- (b) development of a definite/indefinite article contrast from the PPN (and Tongic) specific/non-specific contrast, a change collateral with the development of a distinct article class (from earlier noun markers and number markers).

The principal change to the emerging preposition class of Nuclear Polynesian (apart from loss of PPN *\*aki* 'instrument') was the elimination of PNP *\*a* as an absolutive marker. I would, however, assert that that change is a post-PNP development, and that vestiges of the original function of PPN

\*a are to be found in a number of Nuclear Polynesian languages. For example, *a* is still an optional absolutive marker in Tuvaluan:<sup>14</sup>

TUVALUAN *E kai nee te fafine (a) te ika.*  
 TNS eat ERG ART woman ABS ART fish  
 The woman is eating the fish.

Clark (1976:60) observes that *a* is a plural nominative (absolutive-SPH) in Vaitupu (Tuvalu). That fact can, of course, be restated as the observation that PNP \**a* survives as an absolutive with a  $\emptyset$  plural marker.

The reanalysis of reflexes of PPN \**a* as a proper article (or a plural article in some Samoic/Outlier languages) is simply a generalisation on the residue of functions left for reflexes of \**a* as absolutive (or nominative in nominative-accusative Nuclear Polynesian languages) becomes  $\emptyset$ -marked. (As noted in section 1, I do not regard Chung's (1978) argument that PPN \**i* was an accusative marker as compelling, since it is grounded in facts regarding *nominalisations* of transitive clauses rather than true verbal clauses. I support Clark's original contention that the emergence of reflexes of \**i* as a general accusative marker is a post-PNP innovation.)

The definite and indefinite common noun articles of Nuclear Polynesian, PNP \**te* and PNP \**se*, pose the greatest problem for the reconstruction being proposed here, simply because items reflecting those shapes do not appear in those functions in Tongic. Clark assumes PPN \**sa* 'indefinite' (PTO \**ha*) is irregularly reflected as PNP \**se* 'indefinite'. In the what-you-see-is-what-you-get spirit to which I committed myself in section 4.1, I argue that PNP \**se* 'indefinite' is a reflex of PPN \**se*. The PPN \**sa* reflected as a (non-specific) number marker in Tongic is simply lost in Nuclear Polynesian. PNP \**se* is in fact cognate with PTO \**he* (< PPN \**se*). The function of PPN \**se* is not clear, however. Hitherto I have assumed, by default, that it was the marked-case common noun marker, as in Niuean. But the history of PPN \**se* may have been more complex than that. An interesting possibility is that the original function of \**se* was to introduce (specific indefinite) new referents, and that it became a member of the (specific) common noun marker set in Tongic, but a general indefinite marker in Nuclear Polynesian (displacing the reflex of PPN \**sa*).

PNP \**te* is even more of a problem, since there is no place for a definite common noun article in the PPN noun-marking system proposed here. Clark (1976:65) notes two possible Tongan reflexes of PPN \**te* (ignoring his claims about phonologically irregular reflexes); the first in the pre-pronominal oblique prepositions *'iate/kiate/meiate*, and the second, a fossilised prefix in numerals like TON *teau* 'hundred' and *tekau* 'twenty'. But PNP/PPN \**te* poses greater problems for a comparative-historical linguist committed to the concept of historicity, since clear extra-Polynesian cognates of \**te* are few. It is not obviously reflected in the Fijian languages or in Rotuman, languages that most Oceanists claim to be the nearest relatives of Polynesian. One likely cognate of PPN \**te* is PMC \**te*, reconstructed as a singular common noun article (and as a prefix to unit numerals) in Bender (1981), Harrison (1984) and Harrison and Jackson (1984).

I must confess that I am not sure what to do about PNP/PPN \**te*. Given the standard highest order subgrouping of Polynesian languages, and given that \**te* is *not* irregularly reflected as PTO \**se*/\**e*, the only reason for reconstructing it as a definite article in PPN seems to be that the subgroup of Polynesian in which it is so reflected has more members than the one in which its reflexes are problematic. Since the PPN reconstruction I propose is Tonga-centric, I must favour an analysis in

<sup>14</sup> This was brought to my attention by Niko Besnier.

which the emergence of \*te as a definite article (indeed, the emergence of the *category* definite article) is a Nuclear Polynesian innovation, and that PPN \*te was something other than a definite article – perhaps a number marker, perhaps a pronominal or demonstrative article. Given the evidence, I do not believe that reconstruction to be any less likely than the one proposed by Clark.

#### BIBLIOGRAPHY

- BENDER, B.W., 1981, A fossilized article in Marshallese. In J. Hollyman and A. Pawley, eds *Studies in Pacific languages & cultures in honour of Bruce Biggs*, 209-228. Auckland: Linguistic Society of New Zealand.
- CHUNG, S., 1978, *Case marking and grammatical relations in Polynesian*. Austin: University of Texas Press.
- CHURCHWARD, C.M., 1953, *Tongan grammar*. London: OUP.
- CLARK, R., 1976, *Aspects of proto-Polynesian syntax*. Auckland: Linguistic Society of New Zealand.
- HARRISON, S.P., 1984, Some arguments for pronominal PMC \*te. *Working Papers in Linguistics, University of Hawaii* 16/2:39-61.
- 1989, Lexically-governed preposition choice in Gilbertese. In R. Harlow and R. Hooper, eds *VICALI: Oceanic languages: papers from the Fifth International Conference on Austronesian Linguistics*, 211-236. Auckland: Linguistic Society of New Zealand.
- HARRISON, S.P. and F.H. JACKSON, 1984, Higher numerals in several Micronesian languages. In B.W. Bender, ed. *Studies in Micronesian linguistics*, 61-79. PL, C-80.
- HOHEPA, P., 1969, The accusative-to-ergative drift in Polynesian languages. *Journal of the Polynesian Society* 78:295-329.
- SEITER, W.J., 1980, *Studies in Niuean syntax*. New York: Garland.



PERSONALISED AND NON-PERSONALISED POSSESSION:  
FINAL CONSONANTS IN KUMAK AND OTHER LANGUAGES  
OF FAR NORTHERN NEW CALEDONIA

K.J. HOLLYMAN

1. EXPRESSING POSSESSION

Many New Caledonian languages have two different ways of expressing possession: firstly by suffixing the relevant possessive pronoun (in the discussion to follow this will be represented by the third person singular morpheme common to all the languages involved: *-n*), and secondly by postposing a prepositional phrase, the second element of which is a noun phrase (in the discussion it will be represented by the relevant third person singular personal pronoun). In both methods the grammatical categories of possession and person are inseparable, as indeed the pronominal transform shows.

The semantic categories of noun with which each of these methods is used vary from language to language, as do the declination patterns used with the corresponding nouns. We are not therefore dealing with any expression of a presupposed *mentalité primitive*, nor with the grammatical coding of simple distinctions such as permanent  $\neq$  non-permanent or intimate  $\neq$  non-intimate, but with individually formalised systems of expressing a grammatical relationship.<sup>1</sup>

It is important in this connection to emphasise that some nouns are used with both methods, thereby distinguishing meanings:

KUM *doo-n* 'his spear',            *do-i-ye* 'the spear that hit him';  
CAC *we-n* 'his cord, his vine', *wer-e-i* 'his vein'.

Nouns used with a prepositional phrase to mark possession undergo only predictable phonological change: any final consonant which has become intervocalic because of the vowel serving as preposition is usually replaced by its corresponding intervocalic form: CAC *dek* 'leaf', *deu-n kuic* 'yam leaf', but *dex-e tori* 'taro leaf', where the voiced velar fricative *x* is the predictable intervocalic form of *k*. Possessive suffixation requires more complex and non-predictable change, which is discussed next.

---

<sup>1</sup>For general discussion, see Leenhardt (1946:xxvi-xxix), Haudricourt (1963:9-10).

Robert Blust, ed. *Currents in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, 145-154.  
*Pacific Linguistics*, C-117, 1991.

© K.J. Hollyman



## 2. POSSESSIVE SUFFIXATION: PERSONAL AND NON-PERSONAL POSSESSION

This article studies the suffix method of marking possession as used in the northern New Caledonian languages of Koumac (Fwaxumak or KUM), Pouébo (Caaàc or CAC), Balade-Arama (Yâlayu or YAL) and Oubatche (Jawe or JAW). The first three of these, along with the language of Gomen and Bondé (Yuaanga or YUA), make up the Far Northern Group. The fourth, JAW, with the languages of the Hienghène area (Nemi or NMI, Fwâi or FWA, and Pije or PIJ), belongs to the Northern Group. YUA is quoted essentially to take account of the whole Far Northern Group: in fact it has lost most Proto Oceanic (POC) final consonants, and has not shown the same interest in developing the possessive category as the other languages. JAW on the other hand belongs to the next group southwards, and provides a useful reference both for phonological development and for the use of possessive suffix patterns.<sup>2</sup>

The methods of possessive suffixation divide the nouns into three classes, and reveal the existence of forms marked for non-personalised possession in two of the three. The nouns having a non-personal form as well as the personally possessed forms vary from language to language, showing again the individual and conventionalised nature of the details of the possessive systems.

The role of the non-personalised form overlaps the use of the prepositional phrase to mark possession, since the essential contrast with possessive suffixation, where it occurs, seems in both these other methods to be the distinction of meanings, whether (for example) general as opposed to specific, or 'natural' as distinct from metaphorical. 'General' here implies the name for a 'class' of possessed objects, as opposed to 'specific', 'individual' possession of one or more of them. The examples are chosen simply by their availability in the sources:

CAC *we-n* 'its sap', *we-n kola* 'rainwater', *we-n nu* 'coconut milk', *we* 'water', *we mwaja* 'saliva', *we re kuc* 'sugarcane juice',<sup>3</sup> *we re phii-n* 'his sperm', *we re-rivae-n* 'her tears', *we re thi-n* 'her breast milk'; KUM *wi* 'water', *wi kot* 'rainwater', *wi nu* 'coconut milk', *wi ûûat* 'saliva', *wi idaamaa* 'tears', *wi thi-t* 'breast milk', *wi-t* 'cooking-pot liquid'; YUA *we* 'water, liquid', *we-n* 'his sperm';

KUM *nhu* 'heat, hot', *nhuwa-n* 'his sweat'; CAC *nhu-n* 'his warmth', *nhu le-i* 'his sweat';

KUM *muu-t* 'flower (of plant)', *muu-n* 'its flower', *muu-c* 'wild flower'; CAC *muu-n* 'its flower', *muu-c* 'flowers (as a class)' as in: *neegat kût muuc* 'the sun fades flowers';

CAC *nu-n* 'his coconut palm', *nu re jec* 'forest palms of the *Kentia* section'; KUM *nu* 'coconut palm', *nuu-t* 'the coconut palm in current use'.

## 3. POSSESSIVE SUFFIXATION: NOUN CLASSES

The three classes of nouns with possessive suffixation are now explained in detail.

Class A comprises those never used without a possessive suffix. Their unsuffixed base always ends in -V: KUM *phalavu-n*, YAL *phalaavua-n*, CAC, JAW *phalau-n* '(wife's) brother-in-law,

<sup>2</sup>For the lexical sources, see the Bibliography. The orthography of those sources has been consistently followed.

<sup>3</sup>The preposition *re* means 'in', and noun phrases in which it occurs are not possessive because of its presence. When possession is marked in such a phrase, it is by a suffix attached to the noun following *re*, as in the next three examples quoted in the text. It may also be noted that the degree of unity of these phrases varies: in the phrase for 'tears', *re thivae-n* is phonologically changed to *re-rivae-n*, but this change does not occur to *re phii-n* (x *re-vii-n*) or *re thi n* (x *re-ri-n*).

(husband's) sister-in-law'; KUM, YAL *aa-n*, CAC *â-n* 'garden'. This could be aptly termed 'permanent personal possession', but this must be taken as a simple grammatical description since any semantic implications are obscure.

Class B comprises those with a non-personalised form in -V, to which the possessive suffix is added. There are three subclasses as follows:

Class B1, where -V remains unmodified when the suffix is added: KUM, YAL *mwa*, *mwa-n* 'house'; CAC *nyu*, *nyu-n* 'anchor'.

Class B2, where -V is changed when the suffix is added: CAC *mwa*, *mwe-n*; JAW *nga*, *nge-n* 'house'.

Class B3, where to -V<sub>1</sub> is added V<sub>2</sub> to create the new base: JAW *hwii*, *hwie-n* 'protein food'; CAC (old) *me*, *meu-n*, (modern) *mö*, *möu-n* 'left (hand)'.

Class C comprises those with a non-personalised form (or, more rarely, forms<sup>4</sup>) in -C. The consonants used to mark the non-personal form are not restricted to that purpose, but include the full range of consonants allowed in final position in the given language. In CAC, for example, this means *p*, *t*, *tʃ*,<sup>5</sup> *c*, *k*, *m*, *n*, *ny*, *ng*. The non-personally possessed form is therefore not differentiated from non-possessed nouns by its final consonant, and its special function is in fact validated only by its relation to the system of personal-possessive suffixes. There are five subclasses, the occurrence of which is not predictable:

Class C1, where -C is dropped and the suffix added: KUM *aret*, *are-n* 'bile'; CAC *merip*, *meri-n* 'life, alive'; JAW *wék*, *wé-n* 'cord, vein'.

Class C2, where -C is dropped and the resulting -V changed before the suffix is added: KUM *bwaat*, *bwee-n* 'upper surface'.

Class C3, where -C is replaced by C<sub>2</sub>V before the suffix is added; the replacement pair is not predictable: KUM, YAL *dep*, *devwo-n* 'mat'; CAC *nip*, *nivi-n* 'dream'; KUM *jen*, *jele-n* 'well', *habwan*, *habwali-n* 'blanket', *hiin*, *hiida-n* 'spit'; JAW *javit*, *javile-n* 'reward'.

Class C4, where -C is replaced by V to take the suffix; V is not predictable: CAC *dek*, *deu-n* 'leaf', *hajuk*, *hajui-n* 'true'; CAC, JAW *phac*, *phae-n* 'lungs'; JAW *jic*, *jie-n* 'belly'; CAC *kiic*, *kiie-n*, JAW *kec*, *kei-n* 'liver'.

Class C5, where V is added to -C before the suffix is added; V is unpredictable: CAC, JAW *pun*, *puni-n* 'hair'; CAC *yan*, *yane-n*, JAW *hyan*, *hyane-n* 'contents'; YUA *wal*, *wali-n* 'cord'.

Many of these differences result of course from the different phonological histories of the individual nouns. Thus because \**au* gives a different result in each language, and irrespective of the non-etymological modifications discussed below, POC \**ndau(n)* 'leaf' has five reflexes belonging to four different classes:

<sup>4</sup>This mainly occurs in KUM (for example: *pap*, *pa*, *pavu-t* 'nest': *pavu-n* 'its nest', *pavu-t mââlic* 'bird nest', *pa mabo* 'wasp nest'), but current data are insufficient to determine the different roles of two or more non-personalised forms. It is of some interest to note that the non-personal form may have the equivalent of a prepositional function: see in section 6 reflexes of \**pwatu*.

<sup>5</sup>CAC has many nouns in -*t* which have an alternative form in -*r*, such as *joot/jòdr* 'mountain', *bwat/bwar*, *bwali-n* 'war club', but the alternative does not exist for others, such as *beyit*, *beyi-n* '(any) body joint', *nyagit*, *nyagili-n* 'brain'.

(A) YUA *dòd-n*; (B1) JAW *do, doo-n*; (C1) KUM *duu-t, duu-n*, YAL *daao-t, daao-n*; (C4) CAC *de-k, deu-n*.

This encourages us to check whether, in Class C, the  $-C_1$  of the non-personal form, and any  $-C_2$  of the personally possessed form, are regularly related to etymological consonants.<sup>6</sup> This is now examined.

#### 4. ORIGINS OF NON-PERSONAL POSSESSIVE FORMS

Our springboard is the differential development of medial and final consonants from POC, and its relationship to grammatical categories, as advanced by Françoise Rivierre in her survey of languages of the Hienghène area (i.e. the Northern Group languages). She proposed no explanation for the "epenthetical consonants" which, as she put it, allowed the old final vowel to be maintained (Ozanne-Rivierre 1982:55). It is argued here that for the nouns using possessive suffixation these consonants mark the non-personally possessed form. Examples discussed will be grouped according to the  $-C_1$  of the current non-personal form. The pattern is established through non-possessed nouns, then the possessed are compared with it.

(1)  $-C =$  KUM, YAL, CAC, JAW  $-p$ , YUA  $-\emptyset$ .

\* $-pV$  when  $V$  is lost gives  $-p$  in all except YUA, but when  $V$  is retained, gives  $-v$  in KUM, YAL, CAC, but  $\emptyset$  or  $u$  in JAW and sometimes  $\emptyset$  in CAC: \*Rapi 'evening': CAC *bar/ep*, JAW *bare/ap*; \*upi 'blow (vb)': KUM *uuvi*, YAL *up, uuvet*, CAC *uii*, JAW *üi*.

The pattern is the same in nouns with non-personal form in  $-p$ , and/or the personally possessed in  $-v$  or  $-u$ : \*nipi 'dream': KUM *nivi-n*, YAL *neeve-n*, CAC *nip, nivi-n*, JAW *nep, neu-n*, YUA *noin*; \*ndapu 'ashes': KUM *daap, daavu-n*, YAL *dap*, CAC *dep, devi-n*, JAW *dap*, YUA *da*. There are variations in KUM and CAC with the personally possessed forms of \*maqudip 'life, alive': KUM *malep, maleva-n*, CAC *merip, meri-n*, but note CAC *pha merive-i* 'nourish him'; YAL *maolep*, NMI *matip*, FWA *marip*, PIJ *motip*, YUA *mòdlò*, but with changed  $-C$ , JAW *maric*.

Remodelled to conform in part with this pattern is, for example, \*laya(R) 'sail': KUM *nhap, nhavwu-t, nhavwu-n*, YAL *nhap, nhavwo-t*, CAC *nhep*, JAW *hnep, hneu-n*, YUA *nè*.

Also following these patterns are: 'nest': KUM *papa, pa*, but *pavu-t*, YAL, CAC *mwe/ap*, JAW *nge/ap, ngeau-n*; 'property': CAC *mep, me-n*; 'rope': CAC *mwegep, mwege-n*; 'tooth': CAC *peeyep* (JAW *payoa-n*); 'protein food': CAC *whep, whee-n*, but YAL *wee-n*, JAW *hwii, hwie-n*; 'placenta': KUM *jap, javaa-t*, YAL *jap*.

(2)  $-C =$  KUM, YAL, CAC, JAW  $-t$ , YUA  $-l$ .

(2a) \* $-sV-$ ,  $-jV-$  when  $V$  is lost give  $-t$ , (but  $-l$  in YUA), but when  $V$  is retained give  $-l$  or  $-r$  (depending on which of these the phonemic structure of the language allows): \*nus(*i,o*) 'cuttlefish': KUM, YAL *net*, CAC, JAW *nit*; \*qaso 'sun': KUM, YAL *at*, YUA *al*, CAC *nee/gat*, JAW *hne/gat*.

This pattern is followed by nouns with the non-personal form in  $-t$  and/or personally possessed form in  $-l$  or  $-r$ : \*pajan 'name': KUM *yaara-n*, YUA *yala*, YAL *naat, naara-n*, CAC, JAW *yat*,

<sup>6</sup>The main source of the proto-forms is Ozanne-Rivierre (1982), with some input from Grace (1969), Cashmore (1969), Hollyman (1986b), and from Robert Blust.

*yale-n*; \**qasawa* 'spouse': KUM *aroo-n*, YUA *alò*, YAL *aaroo-n*, CAC *aloo-n*, JAW *kaloo-n*;  
 \**mpuso(s)* 'umbilical cord': KUM, YAL *bora-t*, CAC *bole-n*, JAW *gule-n*; \**susu* 'breast': KUM  
*thi-t*, YAL *the-n*, CAC, YUA *thi-n*, JAW *thi, thi-n*.

Adapted to this pattern are: \**mpuji* 'lock of hair': KUM, YAL *bwili-t*, but CAC *bwili-c*, *bwilie-n*  
 is adapted to pattern (3) below, and JAW *bwili*, *bwilii-n* is regular; \**taji* 'younger sibling': KUM  
*kaari-t*, but YUA *kaali-n*, CAC *celi-n*, JAW *cali-n*.

Following the same pattern for the non-personal form are: \**pu(q)un* 'base, origin': CAC *puu-t*,  
*puu-n*, but JAW *puu-k*, *puu-n* according to pattern (3) below; \**kulit* 'skin': KUM *cii-t*, YAL *ii-t*, but  
 YUA *ci*, CAC, JAW *cii-n*; \**qanunu* 'shadow': KUM *hâlu-t*, YUA *henuul*, but YAL *hâlu-n*, CAC  
*hânu-n*; \**tama* 'father': KUM *kaama-t*, but CAC *cama*, *came-n* and PIJ *tama-n* remain unmodified;  
 \*(*a*)*me(a)* 'tongue': KUM *ku/me-t*, but YUA *ku/mè-n*, YAL *ku/me*, CAC *kuu/me-c*, *kuu/me-n*.

(2b) \**d* has the same result: \**tuqud* 'stand up': KUM *kuut*, YUA *kòdòl*, YAL, JAW *cuut*, CAC  
*coot*.

For possessed nouns we have: \**kudon* 'cooking pot': KUM *ce-t*, *cela-n*, YAL *ye-t*, *yela-n*, but  
 CAC *cet* only, and JAW *cile-k*, *cile-n*; \**tudu* 'knee': KUM *bwaa/jili-t*, YAL *bwaa/jili-t*, but YUA  
*bwa/gili-n*, JAW *ga/jili-n*, CAC *bwa/jire hi-n*.

(2c) *-t* is the most common marker of the non-personalised form in KUM. Other examples include:  
 'paddle, oar': KUM *haa-t*, *haa-n*, YAL *haa-t*, *ha-n*, CAC *haaàt*, JAW *haat*; 'brain': YAL *yagila-t*,  
 CAC *nyagi-t*, *nyagili-n*, YUA *zaagi*; 'war-club': KUM *bwat*, YAL *bwa-t*, *bwa-n*, CAC, JAW *bwa-t*,  
*bwali-n*; and this pattern is followed by many others, among them the borrowing from West Uvean  
*kete* 'basket': KUM, CAC, JAW *keet*, *kee-n*.<sup>7</sup>

(3) *-C* = KUM *-k*, *-c*, YAL *-c*,  $\emptyset$ , YUA  $\emptyset$ , CAC, JAW *-c*, *-k*, in the following correspondences:

	POC	KUM	YAL	YUA	CAC	JAW
(3a)	* <i>t</i>	<i>k</i>	<i>c</i>	$\emptyset$	<i>c</i>	<i>c</i>
(3b)	* <i>q</i>	<i>k</i>	$\emptyset$	$\emptyset$	<i>k</i>	<i>k</i>
(3c)	* <i>k</i>	<i>c</i>	$\emptyset$	$\emptyset$	<i>c</i>	<i>k</i>

(3a) *-C* = YAL, CAC, JAW *-c*, KUM *-k*, YUA  $\emptyset$ .

\**-t* and \**-tV* when V is lost give YAL, CAC, JAW *-c*, KUM *-k*, but when V is retained *-y-* or *-ø-*,  
 and *-x-*, respectively: \**pat* 'four': KUM *-vak*, YAL *-bac*, CAC *-pac*, JAW *-ec*; \**mate* 'dead': KUM  
*maak*, YAL *maac*, CAC, JAW *mac*, YUA *mhá*; \**quta(n)* 'forest': KUM *kak*, YAL, CAC *jec*, JAW  
*kuc*; \**kuRita* 'octopus': KUM *ciixa*, YAL *iiya*, YUA, CAC *ciiya*, JAW *ciia*.

Possessed nouns conforming to this pattern include: \**qate* 'liver': CAC *kiic*, *kiie-n*, JAW *kec*,  
*kei-n*, YUA *kôï-n*, but YAL *kiye-t*; \**natu* 'child': YAL *nae*, *nae-n*, CAC *neie* (*neye*), *nei-n*, JAW  
*naaye*, *naai-n*, the first forms in CAC, JAW deriving from lost etymological *-c* forms (similarly,  
 \**mwata* 'point, tip' is represented only by CAC, JAW *maye* and JAW *mae-n*).

<sup>7</sup>The initial *k-* of POC \**kete* would have given *c-* in KUM, YUA, CAC, JAW, and *y-* in YAL, so the borrowing is  
 clear; cf. Hollyman (1986a:76).

Remodelled to conform are: \*paqa 'thigh': CAC *pha-c*, *phae-n*; \*taliŋa 'ear': CAC *jeeni-c*, *jeeni-n*; \*tian 'belly': CAC *ji-c*, *ji-n*, JAW *ji-c*, *jie-n*.

The ending in *-c* is more popular in CAC, JAW than in KUM, YAL. Also here belong: 'comb': CAC *balic*, *bali-n*, JAW *balic*, *balie-n*; 'embrace': CAC *bovac*, *bova-n*, JAW *goovac*, *goovaa-n*; 'pillow': CAC *bweac*, *bweae-n*, JAW *bweac*, YAL *bweaya*, cf. KUM *bwaak*; 'pulse': CAC *nec*, *neye-n*, YAL *nac*, YUA *nâ*.

(3b) *-C* = KUM, CAC, JAW *-k*, YAL, YUA  $\emptyset$ .

\**-q*, and \**-qV* when V is lost, give KUM, CAC, JAW *-k*, YAL, YUA  $\emptyset$ ; but medial *-q-* gives  $\emptyset$  or more rarely *x*: \**pituqu* 'star': KUM *piguk*, CAC, JAW *piuk*, YAL *piyu*, YUA *pio*; \**tumpuq* 'swelling': CAC *chibuk*, JAW *higuk*; \**tuqud* 'stand up': KUM *kuut*, YAL, JAW *cuut*, CAC *coot*, YUA *kððl*.

This pattern is found with the following possessed nouns: \*(*n*)*dage* 'forehead': CAC *texe-n*, and *bwa/dek*, *bwa/dexe-n*, JAW *ga/dek*, *ga/dexe-n*, KUM *daxa-t*, YAL *daa-n*, YUA *dðð-n*; \*(*n*)*tuqu* 'true': KUM *ke/ma/guk*, CAC *ha/juk*, *ha/jui-n*, JAW *jue/juk*, YAL *hâla/ju*, YUA *mââ/gu*; \**taqu* 'year': CAC *jek*, *jeu-n*, JAW *jak*, YAL *ja*, *jao*, KUM *ja*.

Remodelled on this pattern are: \**kala* 'vulva': CAC, JAW *cane-k*, but KUM, YAL *caalaa-t*, YUA *chê-n*; \**kau* 'penis': CAC *ce-k*, *ceu-n*, but KUM *caau-t*, YUA *chð-n*, JAW *co*.

The *-k* ending is also more common in CAC, and includes: 'middle': CAC *gek*, *geu-n*, but YAL *gaao-t*; 'leg, foot': CAC *hek*, *he-n*, but JAW *he-n*, YAL *kha-n*; 'measure': CAC *jak*, *jaxe-n*, YAL *jak*, KUM *jak* (*jaxe* verb), YUA *jaxè*, cf. JAW *u jak!* 'enough!'; 'nose': CAC *mwadik*, *mwadi-n*, but KUM *mhwadi-t*, JAW *hwâdi-n*, YAL *mh Wade-n*, YUA *mwèèdi-n*.

(3c) *-C* = KUM, CAC *-c*, YAL, YUA  $\emptyset$ , JAW *-k*.

\**-k* and \**-kV* when V is lost, give KUM, CAC *-c*, YAL, YUA  $\emptyset$ , JAW *-k*: \**manuk* 'bird': KUM *mââlic*, CAC *menic*, YAL *mââli*, YUA *mèni*, JAW *manik*; \**namuk* 'mosquito': KUM *nabuc*, CAC *nebwic*, YAL *nabwi*, YUA *néébu*, JAW *naaguk*; but medial *-k-* gives  $\emptyset$  in all five languages: \**pukot* 'net': KUM *piwak*, YAL *puic*, YUA *pwiyo*, CAC *pwiic*, JAW *pwiec*; \**nsoka* 'spear, thorn': KUM, YUA *do*, YAL *do*, *doo-n*, CAC *do*, *do-n*.

Following this pattern are: \**wakaR* 'root': CAC, JAW *waa-n*, YAL *wè-n*, but KUM, YAL *waa-t*; but reduplication leading to aspiration results in a remodelled non-personal form: \**kuku* 'hand, arm': CAC *hi-k*, *hi-n*, but KUM *chi-t*, YAL *yhe-t*.

Adapted to this pattern is: \**kai* 'tree': KUM *cii-c*, CAC *cee-c*, YUA *cee*, JAW *cee-k*, but YAL has *yee* (and not *yee*).

(4) *-C* = *-m* in all five languages.

The only one of these languages with several non-personally possessed nouns in *-m* is CAC, and none of them currently has a POC etymon, unless \**añam* 'weave' is linked with: 'mat': CAC *kâm*, *kâu-n*, JAW *khâm*, cf. YUA *thòm*.

In every other case but one (CAC *tuun*), the nouns show an expanded or modified base for the personally possessed form, and thereby escape any ambiguous interpretation through homophony with forms using the second person singular possessive *-m* 'shadow, shade': CAC *boòðm*,



*boodme-n*, JAW *boom*, YUA *bwòm*, cf. YAL *boom* 'chill'; 'testicles': CAC *jam*, *jabe-n*; 'plate': CAC *tham*, *thabe-n* and *yham*, *yhabe-n*, KUM *shaam*, *shaabe-n*, *shaaba-t*, YAL *ham*, *haba-t*, YUA *yam* and *zam*, *zabo-n*, JAW *helam*, *helame-n*; 'basket': YAL *tholam*, *tholaba-n*; 'fundament': CAC *tuum*, *tuu-n*, cf. JAW *tuun* 'family, clan'.

(5) -C = -n in all five languages except for YUA  $\emptyset$ .

Final POC *n*,  $\eta$  are both lost; but -IV, -ñV, -nV, - $\eta$ V all regularly give -n when -V is lost (and where -l-, -n- survive as intervocalics they give -l- in KUM, YAL, and -n- in CAC, JAW): \*(*kuluk*) 'breadfruit': KUM *cen*, YAL *yen*, CAC, JAW *cin*; \**poñu* 'turtle': KUM, YAL *wan*, CAC *pwin*, YUA *pòn*, *pwòn*, JAW *pwen*; \**qone* 'sand': KUM, YAL *on*, CAC *oon*, YUA *òn*, JAW *kon*; \**tojo* 'mangrove': KUM *kan*, YAL *jan*, CAC, JAW *jen*.

Possessed nouns with this pattern are: \**nsalan* 'road': KUM, YAL *daan*, *dala-n*, YUA *den*, CAC *waa/dan*, JAW *hwaran*; \**qanunu* 'shadow': KUM *hálu-t*, YAL *hálu-n*, CAC *hânu-n*, YUA (irregularly) *henuul*; PEO \*(*ntanu*) 'spirit' (from POC \**qanitu*): YAL *jálu-n*, CAC *jenu-n*, JAW *janu*; \**taliŋa* 'ear': KUM *kiila-t*, YAL *jâli-n*, CAC *jeni-c*, *jeni-n*, YUA *kênii-n*, JAW *jene-n*.

The number of nouns with a non-personal form in -n is small, no doubt because of the possible homophony with the possessed form with third person singular suffix. In any case, all the nouns involved have an augmented base for the personally possessed forms, so that any ambiguity within a given paradigm is avoided.

As with the descendants of \**nsalan* quoted above, the -n of the non-personal form is etymological in: \*(*n*)*tali* 'cord': NMI, FWA, PIJ *dan*, YAL *jan*, *jale-n*; \**pulu* 'hair': CAC, JAW *pun*, *puni-n*, as is its absence in YUA *phu-n*, but -C has been modified in KUM, YAL *pôôlet*.

Following the same pattern are: 'colour, tattooing': KUM, YAL *gaan*, YUA *gaa*, CAC *gan*, *gane-n*, JAW *gan*; 'blanket': KUM *habwan*, *habwali-n*, but YAL *haabwaan*, YUA *habô-n*, JAW *habwen*; 'mark': CAC, JAW *hin*, *hini-n*; 'contents': CAC *yan*, *yane-n*, JAW *hyan*, *hyane-n*, but YUA *hê*, *hê-n*; 'well (waterhole)': KUM *jen*, *jela-n*, but YAL *jem*, CAC *jim*; 'clothes': CAC *yebwin*, *yebwini-n*.

(6) -C = -ng in CAC, YAL, JAW, -ny in KUM, YUA.

The only example is \**wanŋka* 'boat': CAC *wang*, *wage-n*, YAL *waanŋ*, *waanŋa-n*, JAW *wang*, KUM *waan*, YUA *wòny*, *wòjo-n*.

Modified to conform in CAC is: \**qoti* 'end': CAC *ojing*, *oji-n*, cf. KUM *ogi-n*, YUA *ogin*, YAL *oji-n*, JAW *koin*. The other two known cases, and both appear to involve borrowing for some languages, are: 'headgear': CAC *mweeng*, *mweenge-n*, YAL *mweeng*, NMI ('chief's finery') *mweenge-n*, but KUM *mweeng*, YUA *mwéêng*; 'measure': CAC *nhyeng*, *nhyenge-n*, YAL *nhyage*, JAW *hnyeng*, but KUM *nhyage*.

## 5. SUMMARY OF RESULTS

Thus there appears to be a sound basis of normal development from POC showing variation of consonantal change between medial and final positions; and it seems reasonable on the evidence presented to take the view that this differential development was used as the model to extend a system



of personal possessive suffixing, with provision in certain cases for a specific form used for particular semantic values which overrode the personal or individual facet of possession.

So the inherited pattern given by nouns like (1) *\*nipi* (CAC *nip, nivi-n*), (2a) *\*susu* (KUM *thit, thi-n*), *\*ɾajan* (YAL *naat, naara-n*), (2b) *\*kudon* (KUM *cet, cela-n*), (3a) *\*qate* (CAC *kiic, kiie-n*), (3b) *\*(n)dage* (CAC *bwa/dek, bwa/dexe-n*), (3c) (no unmodified possessive example available), was extended to others such as (1) *laya(R)* (KUM *nhap, nhavwu-n*), (2a) *\*mpuji* (KUM *bwilit, bwili-n*), (2b) *\*uRat* (YUA *wal, wali-n*), (3a) *\*paqa* (CAC *phac, phae-n*), (3b) *\*(n)tuqu* (CAC *ha/juk, ha/jui-n*), (3c) *\*kai* (CAC *cee-c*). The systematic modifying was strong enough to integrate a borrowing such as (2c) WUV *kete*.

In adapting more nouns to these inherited patterns, the languages followed their own preferences, giving us the present lack of correspondence of the final consonant marking the non-personally possessed form. The extent of the remodelling also varied with the language concerned, least of all in YUA, most of all in KUM. In CAC, where the largest group of nouns is in Class A, marked for permanent personal possession, the multiplicity of consonants used to mark the non-personally possessed form is the greatest.

Whether or not the non-etymological consonantal marking for non-personalised possession occurred only or mainly where there was no inherited final consonant cannot be affirmed until proto-forms are reconstructed for all the nouns involved. At the moment examples of consonantal substitution (for example *\*maqudip* : JAW *maric*) are very rare. But Class B nouns survived with only vowel marking of non-personal possession.

The ambiguity which could have existed because of the identical shape of inherited final consonants of the nasal series and the singular possessive suffixes (1 KUM, YUA, CAC *-ny*, YAL, JAW *-ng*, 2 all *-m*, 3 all *-n*) has been largely avoided by the use of augmented stems for the personally possessed forms (so that CAC *yebwin* 'clothes' contrasts with *yebwinin* 'his/her clothes'). It may however not be unrelated to the expansion of non-personal forms in *-t* in KUM, YAL, and of those in *-t, -c, -k* in CAC, JAW. Non-personal forms with final nasal consonant which do have homophonous shapes are certainly very rare.

## 6. DIVERSITY OF PATTERNS

When published dictionaries of greater extent than those currently available in printed or typescript form are available, it may perhaps be possible to rediscover under the phonological and morphological complexity some motivating semantic classification of the modes of possession. But it seems certain that it will vary from language to language, as the following examples may serve to show. The sources do not all record the corresponding lexical items, but what they do make available is noted here.

POC *\*pwatu* 'head':

KUM *bwaa-t, bwaa-n* 'head, top', *bwara* 'tip, point', *bwa* 'on, above', *bwaa-t, bwee-n* 'upper part': *bwa kuuvic* 'top of yam', *bwaa mwa* 'roof', *bwaa hoogo* 'mountain top', *bwa wi* 'on the water', *bwara thi-n* 'her nipple', (*dep* 'deck'), *bwa doo* 'ground', *bwee-n* 'on him'.

YUA *bwa, bwa-n* 'on, top', *bwaa-n* 'head': *bwee kwii* 'top of yam', *bwe mwá* 'topmost row of roof thatch', *bwa hōōgo* 'on the mountain, mountain top', (*me thin* 'nipple'), *bwa we* 'on the water'.

YAL *bwa-t*, *bwaa-n* 'head', *bwe* 'on, above', *bwala*, *bwee-t* 'top', *bwala uuvi* 'top of yam', *bwe mwa* 'roof', *bwala juut* 'mountain top', *bwara the-n* 'her nipple', *bwe waang* 'deck', *bwe doo* 'ground'.

CAC *bwa-n* 'head, top', *bwe*, *bwe-n* 'on, surface': *bwana kuic* 'top of yam', *bwana mwá* 'roof', *bwana joot/jòòr* 'mountain top', *bwe we* 'surface of the water', (*guun thi-n* or *maye thi-n* 'her nipple'), *bwe wang* 'deck', *bwana wang* 'prow', *bwana dilic* 'ground'.

JAW *bwa-n* 'head, top', (*xe*, *xe-n* 'on'): *bwana kuic* 'top of yam', *bwana nga* 'roof', *bwana juet* 'mountain top', (*maye wang* 'prow'), *bwana guc* 'ground'.

POC \*uRat 'vein, tendon':

KUM *wat* 'vein, tendon, cord, vine', *wara* + name of horse's body part '(body part) strap', *wara nep* 'reef in sail'.

YAL *wat* 'vein, tendon, cord, vine', *wabili* 'plaited rope'.

YUA *wal* 'vein, vine', *wal*, *wali-n* 'cord'.

CAC *we-t/we-r*, *wer-e-i* 'vein', *we-n* 'cord, vine', *wet o hi-n* 'his nerve', *we bireni* 'plaited rope', *wen ji-n* 'his belt', *wen hi-n* 'his bracelet'.

JAW *wék* 'vein, vine', *wé-k*, *wé-n* 'cord', *wé juwac* 'tendon', *wén jie-n* 'his belt', *wén hi-n* 'his armband', *wén ngen jiget* 'bow-string'.

Kinship terms often figure as examples of 'intimate possession'. Certainly in most of these Far Northern languages, the great majority are marked by pronominal suffix possession. In CAC and JAW they do mainly belong to the group marked by permanent personal possession, but in the other languages there is variation. In KUM, non-personally possessed forms are as good as standard as well, and YAL has them for most kinship nouns. There are therefore no consistent limits to the grammatical expression of family relationships. Names for the basic ones are tabulated to illustrate this point:

	KUM	YAL	YUA	CAC	JAW
father	<i>kââma-t</i>	<i>caama-t</i>	<i>kêê</i>	<i>cama</i>	
	<i>kââma-n</i>	<i>caama-n</i>	<i>kêê-n</i>	<i>came-n</i>	<i>câu-n</i>
mother	<i>axomoo-t</i>	<i>nya-t</i>	<i>nyânyâ</i>		
	<i>axomoo-n</i>	<i>nya-n</i>	<i>nyânyâ i je</i>	<i>ne-n</i>	<i>cê-n</i>
elder sibling	<i>khiiya-t</i>	<i>chiiya-t</i>	<i>pôwa</i>		
	<i>khiiya-n</i>	<i>chiiya-n</i>		<i>chii-n</i>	<i>hie-n</i>
younger sibling	<i>kaari-t</i>				
	<i>kaari-n</i>	<i>caae-n</i>	<i>kaali-n</i>	<i>celi-n</i>	<i>cali-n</i>
first cousin	<i>beega-t</i>		<i>ebiigi</i>		
	<i>beega-n</i>	<i>bega-n</i>		<i>bege-n</i>	<i>haxeli-n</i>
pat.grandfather	<i>kibu-t</i>		<i>wa</i>		
	<i>kibu-n</i>	<i>cebu-n</i>		<i>ciibu-n</i>	<i>ciû-n</i>
grandson	<i>paabuu-t</i>	<i>paabuu-t</i>			
	<i>paabuu-n</i>	<i>paabuu-n</i>	<i>pèèbu-n</i>	<i>pabu-n</i>	<i>paguu-n</i>

## 7. CONCLUSION

It is not claimed that this article explains all the anomalies underlined by George Grace in his article on canonical shapes in KUM. But, along with intra-group and extra-group borrowing, it can account for much that was of concern to him in 1972. Similar concerns for Xârâcùù (Grace 1986) will not be helped, however: the singular possessive suffixes there do not have the form -C, but -CV or -V, and there does not appear to have been a similar development of non-personalised possessed forms in -C. But the history of New Caledonian languages is long enough and idiosyncratic enough (*teste* George!) for the explanations of uncanonical shapes to be both complex and varied.

## BIBLIOGRAPHY

- CASHMORE, C., 1969, Some Proto-Eastern Oceanic reconstructions with reflexes in Southeast Solomon Islands languages. *Oceanic Linguistics* 8/1:1-25.
- CORNE, J.C., n.d., *Projet de dictionnaire de la langue zwanga (ywanga)*. Auckland, Typescript.
- GRACE, G.W., 1969, A Proto-Oceanic finder list. Working Papers in Linguistics, University of Hawaii 1/2:39-84.
- 1972, On canonical shapes in the language of Koumac. In Thomas and Bernot, eds 1972:135-140.
- 1986, Hypotheses about the phonological history of the language of Canala, New Caledonia. In Chris Corne and Andrew Pawley, eds *Le coq et le cagou: essays on French & Pacific languages in honour of Jim Hollyman*, 55-67. *Te Reo* 29.
- HAUDRICOURT, A.G., 1963, *La langue des Nenemas et des Nigoumak: dialectes de Poum et de Koumac, Nouvelle-Calédonie*. Auckland: Linguistic Society of New Zealand.
- 1971, New Caledonia and the Loyalty Islands. In Sebeok, ed. 1971:359-386.
- n.d., *Projet de dictionnaire des langues de l'extrême nord de la Nouvelle-Calédonie (Balade, Arama, Belep)*. Paris. Typescript.
- HAUDRICOURT, A.G. and A. DUFOUR, n.d., *Classement par matières des langues de l'extrême nord de la Nouvelle-Calédonie [bélep, nyâlayu, kumak, caaàc]*. Paris. Typescript.
- HAUDRICOURT, A.G. and F. OZANNE-RIVIERRE, 1982, *Dictionnaire thématique des langues de la région de Hienghène: pi je, fwâi, nemi, jawe*. Paris: SELAF.
- HOLLYMAN, K.J., 1986a, Les emprunts polynésiens dans les langues de la Nouvelle-Calédonie et des îles Loyauté. *Cahiers du LACITO* 1:67-88.
- 1986b, Review of Haudricourt and Ozanne-Rivierre 1982. *Bulletin of the School of Oriental & African Studies* 49/3:619-621.
- n.d., *Dictionnaire caaàc-français*. Auckland. Typescript.
- LEENHARDT, M., 1946, *Langues et dialectes de l'Austro-Mélanésie*. Paris: Institut d'Ethnologie.
- OZANNE-RIVIERRE, F., 1982, Phonologie comparée des langues de Hienghène et du Proto-Océanien. In Haudricourt and Ozanne-Rivierre, 1982: 9-61.
- SEBEOK, T.A., ed., 1971, *Current trends in linguistics*, vol.8: *Linguistics in Oceania*. The Hague Paris: Mouton.
- THOMAS, J.M.C. and L.BERNOT, 1972, *Langues et techniques, nature et société*, vol.1: *Approche linguistique*. Paris: Klincksieck.

## SOME PRACTICAL FACTORS IN THE EARLY DEVELOPMENT OF LINGUISTICS AS A SCIENCE

PATRICIA A. LEE

### 1. INTRODUCTION

This essay shows how some external, practical factors of seventeenth century Europe contributed to what we now think of as the science of linguistics. The factors considered here can be divided into two groups: (1) travel and trade, and (2) printing and literacy. These two factors converged in creating a need for wider communication among peoples and a new, more objective, way of viewing language. In turn these advances led to the creation of universal language plans which contributed to the beginnings of modern phonetics and lexical semantics.

#### 1.1 LINGUISTIC HISTORIOGRAPHY

Intellectual histories can be written from either of two points of view: internal or external. Internal histories trace the ideas and contributions of individual scholars and movements while external histories focus on the external social, cultural and economic forces that influence the course of a discipline. Most historiographers of linguistics concentrate on internal histories, documenting the contributions of specific individuals and tracing the developments of particular schools of thought. (Some notable exceptions to this generalisation are Padley 1976, 1985; Percival 1975, 1986; Robins 1976, 1979, 1988.) A problem with such exclusive attention to individuals and schools is that other contributing factors are, as a consequence, often ignored, and we are left with the impression that discoveries, innovations and trends occurred isolated from any real-world environment. This essay attempts to show how some of the changes that took place in England and France during the seventeenth century affected how language was viewed, and that these changes influenced the subsequent development of scientific linguistics.

#### 1.2 LINGUISTICS AS A SCIENCE

Linguistics as an empirical science is generally considered to have begun in the nineteenth century. In Pedersen's (1931:1) words, "Until the close of the eighteenth century, European linguistic science had advanced but little beyond the knowledge of linguistics achieved by the Greeks and Romans". The immediate reasons for the sudden growth in the development of linguistics as a science were the

---

Robert Blust, ed. *Currents in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, 155-162. *Pacific Linguistics*, C-117, 1991.

© Patricia A. Lee.

rapid expansion of knowledge about newly discovered languages (both Indo-European and non-Indo-European) and the creation of the comparative method. Both these advances and others in the areas of phonetics and semantics can be seen to have been stimulated by the earlier practical concerns stemming from travel and trade on the one hand and printing and literacy on the other. These influences are external in the sense that they are external to the discipline of linguistics. They are practical in that they developed from the social and economic needs of the times, not primarily from intellectual curiosity about the nature of language.

## 2. THE PRACTICAL FACTORS

The practical factors considered here were not, of course, the only external factors affecting the development of linguistics. An enormously important factor which will not be discussed here has been the work of Christian missionaries throughout the world from the Middle Ages to the present. As Robins (1988:475) notes: "...linguistic scholarship marched in step with Christian evangelism, as it has continued to do. The *Propaganda Fide* department of the Roman Church was as prominent in the sixteenth and seventeenth centuries in high-grade linguistic work as the Protestant Summer Institute of Linguistics is today".

### 2.1 TRAVEL AND TRADE

One of the most significant aspects of the seventeenth century was that the known world was expanding as it never had before. Explorers, traders and missionaries were travelling widely for the first time and discovering new peoples, cultures and languages. This had the direct and well-documented effect on linguistics of producing data from a wide variety of languages (see, for example, Pedersen 1931), but it also had another less well-recognised effect. It brought to the attention of scholars the need for communication between speakers of different languages and fostered the idea of creating a new, universal language.

### 2.2 PRINTING AND LITERACY

In the two centuries following the introduction of printing to England in 1476, reading and writing became widespread; this new literacy also contributed to the development of universal languages. In Slaughter's words (1982:10-11):

From all available evidence, it appears that in the middle quarters of the seventeenth century, England became a literate culture as opposed to an oral one. This is not to say merely that some people read or that some people wrote. It is to say the written language and its media and all the social, cultural and cognitive change that goes with it, became institutionalized and predominant. One of the things made possible by printing/literacy was the collection and transmission of vast amounts of information; this required first that there be an adequate language with which to represent that information.

Slaughter also points out that printing had another, cognitive, effect on linguistic scholars of the time in that they began to consider language and its forms in a more objective way. In her view, writing causes language to become decontextualised; it changes language "from an aural, temporal dimension to a visual, spatial one" (Slaughter 1982:40). And this decontextualisation is a crucial step on the way to language becoming an object of scientific study. An important consequence of this

new attitude is that for the first time those concerned with pronunciation were able to clearly make the crucial distinction between letters and the sounds they represented. This innovation is especially prominent in the phonetic sections of Wallis's 1653 grammar (Kemp 1972).

### 3. INTERNATIONAL COMMUNICATION

What we see happening in the seventeenth century is a convergence of two independent kinds of factors in the need, for very practical reasons, of an international language and a rational way of representing that language.

#### 3.1. LATIN AS AN INTERNATIONAL LANGUAGE

Until this time in Western history scholars had used Latin as a universal language of communication, an academic lingua franca. Indeed one of the most significant developments of seventeenth century grammar is a direct consequence of this fact. In 1644 the Port Royal linguist and logician, Lancelot, published his *Nouvelle méthode pour apprendre facilement et en peu de temps la langue latine*; this was the first of many editions of the grammar and was noteworthy for at least two reasons. It was one of the first pedagogical grammars written in the vernacular; but more important in terms of the development of scientific linguistics was its relationship to its companion volumes, *Grammaire générale et raisonnée* (Arnauld and Lancelot 1660) and *La logique, ou, L'art de penser* (Arnauld and Nicole 1662). As Padley (1985:232) points out, "Most authors of vernacular grammars between 1500 and 1700 have no higher aim than the simple – and necessary – production of practical manuals for the classroom or for the use of travellers abroad". *Nouvelle méthode* was no exception to this; it was originally produced for the very practical purpose of teaching Latin as a second language. However, with later editions and the publications of the *Grammaire générale* and *La logique, Nouvelle méthode* (in combination with them) presented some innovative ideas about the nature of language. One such idea was the fundamental distinction between intension (the attributes of what is described by a word) and extension (the things referred to by a word) (Kretzmann 1967:379). Another was the notion that all languages have an underlying, essential, universal character which is a reflection of the rational nature of human beings. Later publications of New Method grammars of Greek, Spanish and Italian exploited this fact. Other ideas that appear in the Port Royal works are that this universal property of languages is logical in nature and that ellipsis is to be explained as (essentially) deletion under identity. This sounds very modern, and indeed Chomsky (1966), in addition to claiming that the Port Royal *Grammaire générale* derives from the philosophy of Descartes, has suggested that it is the real precursor to transformational grammar.

Regardless of the controversy surrounding the Cartesian nature of these works (Aarsleff 1970; Lakoff 1969; Salmon 1969), one of Chomsky's points is valid. The Port Royal grammarians considered language in a way very similar to modern generative grammarians, with abstract, underlying forms and transformational-type rules for deriving surface structures. These scholars can be seen as making a transition from the essentially practical endeavour of pedagogy to the beginnings of the scientific approach to syntax and semantics. Their idea that there is an underlying, rational, universal level common to all languages was adopted by a generation of language planners who attempted to put this theoretical tenet to practical use in the development of an artificial, universal language to replace Latin as the international language for travellers and scholars alike.



### 3.2 UNIVERSAL LANGUAGE SCHEMES AND SEMANTICS

Despite the attempts of scholars and educators like those at Port Royal, Latin was losing its position as a practical universal language for several reasons. For one thing, not everyone who needed a lingua franca was a scholar; the new merchant class needed to be able to communicate with speakers of other languages, but had neither time nor inclination to spend hours in the classroom learning Latin. For another thing, Latin had been closely associated with the Catholic Church which, due to the reforms of the period, no longer maintained the widespread prestige (or even approval) that it had earlier. And finally, even among those who had learned Latin as children, there were at this time so many different pronunciations (influenced, naturally enough, by the speakers' native languages) that it was no longer entirely mutually intelligible across speakers.

This need for a new international language, along with the growing attention to a scientific approach to matters and the increasing number of literate people, led to the development of artificial languages or systems of 'real characters'. These invented, 'auxiliary' languages were to be independent of any known natural language, were based on rational organisation, and consisted of taxonomies with associated writing systems that were intended as accurate representations of all the ideas and things in the known world. The culmination of these language plans, Wilkins's *Essay towards a real character* (1668), has been aptly referred to as 'a grammar of things' (Cohen 1977:30).

Other prominent language planners were Dalgarno and Leibniz and it has been noted that some of their work was motivated by more than one kind of pragmatic factor: "It is clear that Leibniz intended the *caractéristique universelle* to supply not only a clear, regular, unequivocal means of communicating between scholars, but also an instrument of reasoning that, if composed with the greatest care, would, he believed, lead inevitably to a certainty of conclusion comparable with that found in mathematical demonstrations" (Knowlson 1975:109).

The essentials of the universal languages were that their words were organised into a rational system, that they were built up from a set of basic elements or characters ('semantic primes' in modern parlance), and that the written forms of the words reflected their meanings. It is often noted that this idea was a direct consequence of the discovery of Chinese and its logographic writing system. (For more details on these language schemes, see Breckle 1975; Cohen 1977; Knowlson 1975; Large 1985; Salmon 1972, 1979, 1986; Slaughter 1982.)

As is well documented in Slaughter, the universal language schemes were closely allied to the development of taxonomical science; one of the major concerns of the time was nomenclature: how to systematically refer to all of the new things being discovered by travellers and explorers. This taxonomising of knowledge helped form the foundation for modern semantics.

Modern synchronic semantics can be conveniently divided into lexical and logical varieties. Except for the foreshadowing by the Port Royalists mentioned above, logical semantics is essentially a product of the twentieth century. Lexical semantics, on the other hand, is based on the notions of taxonomy and relations very like those that the language planners of the seventeenth century devised. Wilkins's *Essay* (1668) was a source of inspiration to one of the nineteenth century's great lexicological efforts, *Roget's Thesaurus* (1852), which in turn exerted an influence on subsequent linguists, including Bally (Firth 1937:74; Salmon 1979:202-203).

In more contemporary terms, Dolezal (1987:271) shows how Wilkins, in constructing his 'philosophical tables' uses the concept of semantic primitives and hierarchy which "can be seen as a development of a structural semantics". More specifically, Dolezal (p.280) argues that the method

that Wilkins used was essentially the same as modern componential analysis in that he made use of techniques that are very similar to Nida's criterial processes of "naming, paraphrasing, defining and classifying". In addition, Wilkins's hierarchical classification uses the basic sense relations (in modern terminology) of synonymy, antonymy and hyponymy which are central to modern lexical semantics.

Universal writing systems, which evolved from an earlier concern with cryptography and stenography, were an important component of most of the seventeenth century artificial languages (Cohen 1977:13-17; Knowlson 1975:18-20). In the late sixteenth century interest developed in creating secret writing systems in which symbols represented whole words (as in Chinese) or, in some cases, whole sentences. This efficient way of coding language was soon adopted by Bright (1588, 1628) as a means of recording public linguistic events, such as religious sermons and court proceedings. Bright's symbols were isomorphic to words and formed the inspiration behind the notion that an ideal universal language should have an associated universal writing in which each symbol represented one concept.

Other aspects of writing were also of importance to the scholars of this period; Dalgarno, Holder and Wallis, for example, were concerned with teaching writing to deaf mutes, and the beginning of phonetics in Britain was closely allied with attempts to rationalise and standardise English spelling.

#### 4. PHONETICS IN ENGLAND

The two most crucial factors in the early development of scientific phonetics in the West were writing systems for universal language and the reformation of spelling. Among those language planners whose interest in writing led them to a concern with phonetics were Lodwick and Wilkins. Lodwick (1647) is generally credited with devising an alphabetic system which displays natural phonetic classes not unlike today's generative phonologists (Abercrombie 1948; Fromkin and Ladefoged 1981; Salmon 1972). Wilkins, in order to devise a writing system as closely related to pronunciation as possible, carefully investigated how each sound was made, and organised all the observed sounds into charts based essentially on place and manner of articulation. He also used binary features, discussed suprasegmentals, and used minimal pairs in his analyses (Subbiondo 1987). In Wilkins's *Essay* we find drawings of articulations for vowels and consonants that are surprisingly modern.

The other major impetus behind the development of phonetics was the standardisation of spelling, and this was not (as it is sometimes assumed) the work of Caxton and subsequent printers. Brengelman (1980) convincingly argues that the regularisation of spelling was a concerted effort of the seventeenth century linguistic scholars, including not only Lodwick and Wilkins but also Wallis and Holder.

John Wallis could justifiably be called the father of articulatory phonetics. He was the first to systematise phonetic descriptions in terms of place and manner of articulation and to propose a universal theory of phonetics based upon it (Constantinescu 1974; Fromkin and Ladefoged 1981). The sophisticated system of phonetics found in Wilkins's work is based upon Wallis's descriptions. Kemp (1972:40-66) points out that Wallis introduced innovations in the descriptions not only of segmental features but also of pitch, sentence intonation and syllables. Wallis is also the first to consistently succeed in distinguishing letters from sounds.

Holder's later work (1669) builds upon Lodwick's features and includes distinctions of voicing and obstruency; he also presents "phonetic properties which serve to specify all possible sounds and classes of sounds" (Fromkin and Ladefoged 1981:4). The combined efforts of these four orthoepists and lexicographers provided the foundations of the science of phonetics.

## 5. THE BEGINNINGS OF COMPARATIVE LINGUISTICS

We saw above that Leibniz was one of the foremost of the universal language planners; he also played a role in another advance of the period. Although his main ideas about language families turned out to be wrong, based as they were on 'some sort of inspired intuition' rather than rigorous analysis of data, he did help to establish the systematic collection and documentation of the world's languages (Pedersen 1931:9). He, along with kindred souls, including Catherine II of Russia, encouraged traders, explorers and missionaries to catalogue the languages they encountered in their travels. This effort culminated in Adelung's classic *Mithridates* (1806 – 1817), which helped prepare the foundations for modern historical-comparative linguistics.

## 6. CONCLUSION

The importance of the seventeenth century Port Royal *Grammaire générale* to grammatical theory has long been recognised. And it is almost tautological to say that the science of comparative linguistics could not have developed without the discoveries made about the languages of Europe, Asia, Africa and America in the seventeenth and eighteenth centuries. What is less obvious is that during the seventeenth century these discoveries, in conjunction with the spread of literacy, formed the foundations for modern phonetics and lexical semantics.

A cross-linguistic need for communication arose in this period because of the encounters that travellers and traders had with speakers of new languages. An important consequence of this was the development of universal language plans, which included the fundamental sort of taxonomies that modern lexical semantics rests upon, and which encouraged an interest in the relationship between writing and speech. The growth of the printing industry and the resulting widespread literacy led to attempts to regularise spelling and orthography which, in conjunction with universal writing, resulted in the beginnings of scientific phonetics.

## BIBLIOGRAPHY

- AARSLEFF, Hans, 1970, The history of linguistics and Professor Chomsky. *Language* 46:570-585.
- AARSLEFF, Hans, Louis G. KELLY and Hans-Josef NIEDEREHE, eds, 1987, *Papers in the history of linguistics*. Amsterdam/Philadelphia: John Benjamins.
- ABERCROMBIE, David, 1948, Forgotten phoneticians. *Transactions of the Philological Society* 1-34.
- ADELUNG, J., 1806 – 1817, *Mithridates oder allgemeine Sprachenkunde*. 4 vols. Berlin.
- ARNAULD, Antoine and Claude LANCELOT, 1660, *Grammaire générale et raisonnée, contenant les fondemens de l'art de parler, expliquée d'une manière claire et naturelle*. Paris. (1966 edn).
- ARNAULD, Antoine and Pierre NICOLE, 1662, *La logique, ou, L'art de penser...* Hildesheim: G. Olms; Paris: Charles.
- BRECKLE, Herbert E., 1975, The seventeenth century. In Thomas A. Sebeok, ed. *Current trends in linguistics*, vol. 13: *Historiography of linguistics*, 277-382. The Hague/Paris: Mouton.

- BRENGELMAN, F.H., 1980, Orthoepists, printers and the rationalization of English spelling. *Journal of English and Germanic Philology* 79:332-354.
- BRIGHT, Timothy, 1588, *Characterie: an arte of shorte, swifte, and secrete writing by character*. London.
- BRIGHT, Timothy, 1628, *The art of stenography*. London.
- BYNON, Theodora and F.R. PALMER, eds, 1986, *Studies in the history of Western linguistics: in honour of R.H. Robins*. Cambridge: University Press.
- CHOMSKY, Noam, 1966, *Cartesian linguistics*. New York: Harper.
- COHEN, Murray, 1977, *Sensible words: linguistic practice in England, 1640-1785*. Baltimore/London: The Johns Hopkins University Press.
- CONSTANTINESCU, Ilinca, 1974, John Wallis (1616-1703): a reappraisal of his contribution to the study of English. *Historiographia Linguistica* 1/3:297-311.
- DALGARNO, George, 1661, *Ars signorum*. London: J. Hayes.
- DOLEZAL, Frederic T., 1987, John Wilkins and the development of a structural semantics. In Aarsleff, Kelly and Niederehe, eds, 1987:271-281.
- FIRTH, J.R., 1937, *The tongues of men*. London: Watts.
- 1946, The English school of phonetics. *Transactions of the Philological Society*, 92-132.
- FROMKIN, Victoria A. and P. LADEFOGED, 1981, Early views of distinctive features. In R.E. Asher and Eugénie J.A. Henderson, eds *Towards a history of phonetics*, 3-8. Edinburgh: University Press.
- HOLDER, William, 1669, *Elements of speech: an essay of inquiry into the natural production of letters, with an appendix concerning persons deaf and dumb*. London: Martyn.
- KEMP, J.A., 1972, *John Wallis' Grammar of the English language*. London: Longman.
- KNOWLSON, James, 1975, *Universal language schemes in England and France 1600-1800*. Toronto: University of Toronto Press.
- KRETZMANN, Norman, 1967, History of semantics. In Paul Edwards, ed. *The encyclopedia of philosophy*, 358-406. New York: Macmillan.
- LAKOFF, Robin, 1969, Review of A. Arnauld and C. Lancelot, *Grammaire générale et raisonnée*, Paris, 1660 (1966 edn). *Language* 45:343-364.
- LANCELOT, Claude, 1644, *Nouvelle méthode pour apprendre facilement et en peu de temps la langue latine*. Paris.
- LARGE, Andrew, 1985, *The artificial language movement*. London: Blackwell.
- LEIBNIZ, Gottfried Wilhelm, 1666, *De arte combinatoria*. Leipzig.
- LODWICK, Francis, 1647, *A common writing*. London.
- PADLEY, G.A., 1976, *Grammatical theory in Western Europe 1500-1700: The Latin tradition*. Cambridge: University Press.
- 1985, *Grammatical theory in Western Europe 1500-1700: trends in vernacular grammar I*. Cambridge: University Press.
- PEDERSEN, Holger, 1931, *The discovery of language: linguistic science in the 19th century*. Bloomington: Indiana University Press.
- PERCIVAL, W. Keith, 1975, Grammatical tradition and the rise of the vernaculars. In Thomas A. Sebeok, ed. *Current trends in linguistics*, vol. 13: *Historiography of linguistics*, 231-275. The Hague/Paris: Mouton.
- 1986, Renaissance linguistics: the old and the new. In Bynon and Palmer, eds 1986, 56-68.
- ROBINS, R.H., 1976, Some continuities and discontinuities in the history of linguistics. In Herman Parret, ed. *History of linguistic thought and contemporary linguistics* 13-31. New York: de Gruyter.
- 1979, *A short history of linguistics*. London: Longman.
- 1988, Appendix. History of linguistics. In Frederick J. Newmeyer, ed. *Linguistics: The Cambridge survey*, vol. 1: *Linguistic theory: foundations*, 462-484. Cambridge: University Press.
- SALMON, Viviau, 1969, Review of Noam Chomsky *Cartesian linguistics*, New York, 1966. *Journal of Linguistics* 5:165-187.

- 1972, *The works of Francis Lodwick: a study of his writings in the intellectual context of the seventeenth century*. London: Longman.
- 1979, *The study of language in 17th century England*. Amsterdam: John Benjamins.
- 1986, Effort and achievement in seventeenth-century British linguistics. In Bynon and Palmer, eds 1986, 69-95.
- SLAUGHTER, M.M., 1982, *Universal languages and scientific taxonomy in the seventeenth century*. Cambridge: University Press.
- SUBBIONDO, Joseph L., 1987, John Wilkins' theory of articulatory phonetics. In Aarsleff, Kelly and Niederehe, 1987, 263-270.
- WALLIS, John, 1653, *Grammatica linguae Anglicanae*. Oxford.
- WILKINS, John, 1668, *An essay towards a real character, and a philosophical language*. Oxford.

## VOWEL DELETION AND VOWEL ASSIMILATION IN SEDIQ

PAUL JEN-KUEI LI

In this paper I shall reformulate some phonological rules and discuss some special problems concerning the vowel-deletion and vowel-assimilation rules in Sediq, an Austronesian language spoken in Taiwan.<sup>1</sup>

In the Paran dialect of Sediq, the following phonological rules have been worked out in order to derive most of the surface phonetic forms of its verbs (see Yang 1976; Li 1977):<sup>2</sup>

(1a)  $\left\{ \begin{array}{c} p \\ b \end{array} \right\} \rightarrow k/ \_ \#$

Imperative	Stem	Base	
<i>tap-i</i>	<i>atak</i>	/atap/	to cut (with scissors)
<i>yup-i</i>	<i>piyuk</i>	/piyup/	to blow (breath)
<i>leb-i</i>	<i>eluk</i>	/eleb/	to close
<i>ruburub-i</i>	<i>ruberuk</i>	/rVberub <sup>β</sup> /	to broil

(1b)  $m \rightarrow \eta/ \_ \#$

<i>gulem-i</i>	<i>geeluj</i>	/gVelem/	to bend
<i>tulam-i</i>	<i>talaŋ</i>	/talam/	to run

<sup>1</sup>Professor George Grace first introduced me to the field of Austronesian linguistics when I was one of his graduate students at the University of Hawaii from 1967 to 1970 and from 1972 to 1973. I could not have successfully completed my studies there without his encouragement, patience, and efforts to secure a teaching assistantship for me. I am indebted to Robert Blust, Zhong-wei Shen, Pang-hsin Ting and Hsiu-fang Yang for comments on earlier versions of this paper. I have also profited from discussion with William S.-Y. Wang.

<sup>2</sup>Hsiu-fang Yang (1976) gives a detailed description and a fairly thorough analysis of Sediq phonology. She has worked out most of the phonological rules, but has overlooked a few others. (For instance, she made no mention of the palatalisation rule  $t \rightarrow c$  and  $d \rightarrow j$  before  $i$  or  $y$  in the speech of younger speakers in some Sediq dialects; see Li 1982.) Based on my own observations, I have reformulated or reinterpreted some of her rules, especially rules 4, 5, 7 and 8 in this paper. The rest of the rules illustrated here provide a necessary background for understanding the problems I shall discuss herein.

<sup>3</sup>The symbol 'V' stands for a vowel that is indeterminable in its underlying form.



$$(2) \left\{ \begin{array}{c} t \\ d \end{array} \right\} \rightarrow cl\_ \#$$

<i>kuret-i</i>	<i>keruc</i>	/keret/	to cut
<i>quyut-i</i>	<i>qiyuc</i>	/qiyut/	to bite
<i>rutud-i</i>	<i>rutuc</i>	/rutud/	to link
<i>huged-i</i>	<i>haguc</i>	/haged/	to cook

$$(3) l \rightarrow n/\_ \#^4$$

<i>tubul-i</i>	<i>tabun</i>	/tabul/	to dig
<i>tukul-i</i>	<i>takun</i>	/takul/	to fall

$$(4) \left\{ \begin{array}{c} e \\ o \end{array} \right\} \rightarrow w/\_ (C) \#^5$$

Cf. Yang's rule:  $\left\{ \begin{array}{c} e \\ o \end{array} \right\} \rightarrow w/\_ (C) \#$

<i>deheq-i</i>	<i>dehuq</i>	/deheq/	to arrive
<i>subet-i</i>	<i>sebuq</i>	/sebet/	to hit, to thresh
<i>toos-i</i>	<i>taus</i>	/taos/	to beckon
<i>tuduroy-i</i>	<i>tudoruy</i>	/tVdoroy/	to roll down

$$(5a) g \rightarrow w/\left\{ \begin{array}{c} u \\ a \end{array} \right\} \_ \#^6$$

<i>kucug-i</i>	<i>kiicuw</i>	/kVicug/	to fear
<i>luhug-i</i>	<i>luhuw</i>	/luhug/	to thread
<i>rgag-i</i>	<i>*rigaw &gt; rigo</i>	/rigag/	to fool around
<i>sapag-i</i>	<i>*sapaw &gt; sapo</i>	/sapag/	to lay mats

Synchronically both /g/ and /aw/ are derived as /o/[ɔ:] in the Paran dialect of Sediq (see (6a) below). However, our diachronic evidence indicates that the process of sound change was: \*ag > aw > o. Some Atayal dialects such as Mayrinax and Matabalay still retain the distinction between word-final /g/ and /aw/ (see Li 1981). There is also some synchronic evidence for the two step derivation. A generalisation can be gained by assuming that /g/ is derived as /w/ after a back vowel, not only after /u/, but also after /a/ (see examples in (5a)).

<sup>4</sup>This rule applies only to the speech of younger speakers.

<sup>5</sup>The phonetic environment given in Yang's (1976) rule is after stress. However, our comparative evidence indicates that it is the vowel in the final syllable, especially *e* < \**a*, that becomes *u* in all Atayalic languages and dialects (see Li 1981). There are only a few examples for the change *o* > *u* in our data, and there is no compelling evidence to show that the phonetic environment refers to stress or to the final syllable. However, the same environment can be assumed for the change of both mid-vowels to capture a generalisation if the rule *o* > *u* is valid.

<sup>6</sup>I do not believe that the semivowels /w, y/, which derive from /g/, are deleted in word-final position, as stated in Yang (1976:654-655). As a matter of fact, the loss of /g/ in word-final position produces a compensatory lengthening of the preceding vowel, which is interpretable as /w/ or /y/ depending on the preceding vowel, for example, *kucug-i*, *kiicuw* 'to fear', *brig-i*, *baruy* 'to buy'.

(5b)	<i>g</i>	--->	<i>y/i_#</i> <sup>7</sup>		
	Stem		Imperative	Base	
	<i>brig-i</i>		<i>baruy</i>	/barig/	to buy
	<i>pcig-i</i>		<i>pacuy</i>	/pacig/	to carve
	<i>rugurig-i</i>		<i>rugeruy</i>	/rVgerig/	to sift
(6a)	<i>aw</i>	--->	<i>o/_#</i>		
	<i>sunaw-i</i>		<i>sino</i>	/sinaw/	to wash (clothes)
(6b)	<i>ay</i>	--->	<i>e/_#</i>		
	<i>ruḡay-i</i>		<i>raḡe</i>	/raḡay/	to play
	<i>kusuḡay-i</i>		<i>kusaḡe</i>	/kVsaḡay/	to rest

VOWEL DELETION BEFORE STRESS

- (7a) V --->  $\emptyset/_C^n \acute{V}$  where C must be a true consonant<sup>8</sup>  
 Cf. Yang's rules:  $\acute{V}$  ---> *u/\_(C)\acute{V}*  
*u* --->  $\emptyset/\#\_C\acute{V}$

Stress mostly falls on the penultimate syllable of the word in all Sediq dialects. All word-initial vowels get deleted, as illustrated in (7b). Any other vowel before a true consonant that is immediately followed by the stressed vowel, and all vowels between consonants before a stressed syllable are reduced to a phonetic [u] or [ə], depending on the dialect or speaker, as in (7c).

	Stem	Imperative	Base	
(7b)	<i>atak</i>	<i>tap-i</i>	/atap/	to cut (with scissors)
	<i>imah</i>	<i>mah-i</i>	/imah/	to drink
	<i>usa</i>	<i>sa-i</i>	/usa/	to go
	<i>eyah</i>	<i>yah-i</i>	/eyah/	to come
	<i>obuh</i>	<i>buh-i</i>	/obuh/	to bake
(7c)	<i>kapah</i>	<i>kupah-i</i>	/kapah/	to adhere
	<i>qita</i>	<i>quta-i</i>	/qita/	to see
	<i>pu-suwak</i>	<i>pu-suwak-i</i>	/pV-suwak/	to yawn <sup>9</sup>
	<i>betaq</i>	<i>butaq-i</i>	/baytaq/ <sup>10</sup>	to stab
	<i>tudoruy</i>	<i>tuduroy-i</i>	/tVdoroy/	to roll down

It is clear from the above data that Sediq has a general vowel-deletion rule. Only stressed vowels retain their full vowel value, whereas vowels before stress either get deleted or reduced to a phonetic [u] (or [ə] in the other dialects). Since the reduced vowel is phonetically predictable, it can be left out

<sup>7</sup>It is not entirely certain whether /g/ after /i/ surfaces as [y]. The fact that the back vowel /u/, instead of the expected front vowel /i/, turns up in the final syllable of the stem, for example, *baruy*, presents a problem.

<sup>8</sup>True consonants are consonants produced by constriction or closure of the oral or nasal passage. Semivowels are also true consonants, but pharyngeal and glottal consonants are not (see below).

<sup>9</sup>The full vowel /u/ in the stem/suwak/ is reduced to a weak phonetic [u] or [ə] before stress.

<sup>10</sup>Both internal and comparative evidence indicates that the diphthongs *ay* and *aw* are realised as *e* and *o* respectively in non-final syllables (see Li 1981:270-272).

in a broad phonetic transcription. All that is needed is a late phonetic rule to insert a phonetic [u] (or [ə]) between certain consonants (see Li 1977).

As we shall see in the following assimilation rule (8a), a vowel in an ante-penultimate syllable in certain phonetic environments may not get deleted or reduced.

## VOWEL-ASSIMILATION RULE

$$(8a) \quad V \quad \text{---} \rightarrow V_x/C\_ \left\{ \begin{array}{l} h \\ \text{ʔ} \end{array} \right\} \acute{V}_x$$

Cf. Yang's rule:  $u \text{ ---} \rightarrow V_x/C\_ (h)\acute{V}_x$

A vowel in the ante-penultimate syllable gets assimilated to the immediately following stressed vowel. There may be an intervening pharyngeal fricative /h/ or glottal stop /ʔ/, but no true consonants. As stated in footnote 8, true consonants are produced in the oral or nasal passage. In the forms in the second column in (8b) below, the vowel in the ante-penultimate syllable is identical with that in the penultimate, whereas there is no such assimilation in the forms in the first column. As stated in (7a) above, all vowels before stress in the other phonetic environments get deleted or reduced.<sup>11</sup>

(8b)	<i>suliq-i</i>	<i>seelaq</i>	/sVelaq/	to kill (an animal)
	<i>kucug-i</i>	<i>kiicuw</i>	/kVicug/	to fear
	<i>suluhe</i>	<i>sulahay-i</i>	/sVluhay/	to learn
	<i>dehuq</i>	<i>deheq-i</i>	/deheq/	to arrive
	<i>tuleʔuŋ</i>	<i>tuluʔuŋ-i</i>	/tVleʔuŋ/	to sit down
(8c)	<i>qumiyuc</i>		/q-um-iyut	to bite
	<i>tukan-i</i>		/tekan-i/	Pound!
	<i>sukusik-i</i>		/sikesik-i/	Sweep!
	<i>subugihur</i>		/sVbVgihur/	to blow (wind)
	<i>mu-cehaq</i>		/mu-cehaq/	to copulate
	<i>mu-burinah</i>		/mu-bVrinah/	to return
	<i>luwah-i</i>		/lawah-i/	Open!
(8d)	<i>reus</i>	to bury		
	<i>taus</i>	to beckon		
	<i>dehuq</i>	to arrive		
	<i>boh-i</i>	Bake!		

The data in (8c) show that only the true consonants will block the vowel-assimilation rule. Unlike the forms in the second column in (8b), the vowel in the ante-penultimate syllable is not assimilated to the following vowel in (8c). The disyllabic forms in (8d) are excluded by the rule since assimilation applies only to an ante-penultimate vowel.

Let us examine the following verb inflections that present a problem, noted but not resolved by Yang (1976:663):

<sup>11</sup>Headings are not provided for the following rules because (8b)-(8d) contain both the stem and the imperative suffix *-i*. In addition, (8c) contains the future prefix *mu-*.

	Present	Future	Stem	Imperative	Others
(9a)	<i>r-um-ehak</i>	<i>mu-rehak</i>	<i>rehak</i>	<i>rehēp-i</i>	to sow
	<i>c-um-ehak</i>	<i>mu-cehak</i>	<i>cehak</i>	<i>cehēp-i</i>	<i>cehēp-un</i> to lick
	<i>r-um-ehaq</i>	<i>mu-rehaq</i>	<i>rehaq</i>	<i>reheq-i</i>	<i>reheq-un</i> to remove
					<i>reheq-an</i>
	<i>mu-kuseaŋ</i>	<i>mu-pu-kuseaŋ</i>	<i>kuseaŋ</i>	<i>kuseeŋ-i</i>	<i>kuseeŋ-un</i> to quarrel

The four verbs above are the only ones that we have found with these peculiar inflections. Notice that the second vowel in the verb stem is *a* when it appears in the final unstressed syllable, but *e* when it occurs in the penultimate stressed syllable followed by a suffix *-i*, *-un* or *-an*. If *e* is treated as the base, as we have done with all the other vowels, then *a* will be derived as a phonetic output in the final syllable:

(9b)  $e \rightarrow a/_{(C)\#}$

We shall then find many counter-examples such as the following, which do not surface as *a* in the final syllable:

	Present	Future	Stem	Imperative	Passive
(10)	<i>k-um-barux</i>	<i>mu-kubarux</i>	<i>kubarux</i>	<i>kubarex-i</i>	to borrow
	<i>m-eluk</i>	<i>me-eluk</i>	<i>eluk</i>	<i>leb-i</i>	<i>leb-un</i> to close
	<i>q-m-epu</i>	<i>mu-qepu</i>	<i>qepu</i>	<i>qupe-i</i>	<i>qupe-un</i> to filter
	<i>mu-pu-dehug</i>		<i>dehuq</i>	<i>deheq-i</i>	to arrive

In the above verbs, the vowel *e* in the second syllable of the stem does not become *a* when it appears in the word-final syllable as in (9). See also rule (4) above: in fact,  $e \rightarrow u/_{(C)\#}$  is a very general rule in the language.

Alternatively, if we treat *a* in the second syllable of the stem as the base and *e* in the suffixed forms as derived, and attribute it to the process of assimilation with the preceding vowel, then there will be no exceptions. The main problem with this solution is: how can we account for the fact that the stressed vowel gets assimilated to the preceding unstressed vowel only in these forms? In all the other cases, the stressed vowel retains its full vowel value and is not affected by its adjacent segments. Moreover, all other instances show that it is the vowel in the ante-penultimate syllable that gets assimilated to the following stressed vowel. For example, when the prefix *mu-* 'future' is attached to a disyllabic verb stem beginning with a vowel, the vowel /*u*/ in the prefix becomes assimilated to the following vowel of the stem, as illustrated in Yang (1976):

(11)	/mu-eyah/	$\rightarrow$	<i>me-eyah</i>	to come
	/mu-adis/	$\rightarrow$	<i>ma-adis</i>	to bring
	/mu-imah/	$\rightarrow$	<i>mi-imah</i>	to drink
	/mu-ekan/	$\rightarrow$	<i>me-ekan</i>	to eat
	/mu-aŋal/	$\rightarrow$	<i>ma-aŋal</i>	to take

Shall we, then, treat the rule  $a \rightarrow e$  for the four verbs in (9) as a minor rule, which applies only to a few verbs? A common feature among these verbs is that the vowel in the penultimate syllable is [e], historically derived from \**a*, which might not have received any stress at an earlier stage. Stress might have fallen on the final syllable with the vowel [a] < \**a* in the proto-forms. As it shifted its position when a suffix was attached to it, it was susceptible to change, that is, it was reduced to a schwa or assimilated to the preceding vowel which was a schwa. The stress, therefore, still fell on

the final syllable, that is, the suffix in the proto-forms. Such proto-forms (with stress on the final syllable) may have been rare.

The following examples seem to be exceptions to the vowel-deletion rule. Instead of being reduced to [u] as expected, the vowel /i/ in the ante-penultimate syllable seems to have assimilated to the immediately following /y/:<sup>12</sup>

	Present	Future	Stem	Imperative	
(12)	<i>d-um-ayo</i>	<i>mu-dayo</i>	<i>dayo</i>	<i>diyag-i</i>	to help
	<i>k-um-ayak</i>	<i>mu-kayak</i>	<i>kayak</i>	<i>kiyap-i, kiyap-un</i>	to cut
	<i>mu<sup>?</sup>uyas</i>		<i>pu<sup>?</sup>uyas</i>	<i>piyas-i</i>	to sing
	<i>m-hiyuw</i>	<i>mi-hiyuw</i>	<i>hiyuw</i>	<i>hiyeg-i</i>	to stand
	<i>puriyux</i>	<i>mu-puriyux</i>	<i>puriyux</i>	<i>puriyux-i</i>	to change
	<i>c-um-iyuk</i>	<i>mu-ciyuk</i>	<i>ciyuk</i>	<i>ciyuk-i</i>	to answer

Then we are confronted with the problem of how to reformulate the assimilation rule. Should /y/ and /i/ be assigned almost the same features?

The vowel-deletion and vowel-assimilation rules apply in different phonetic environments. Since they are in complementary distribution, the rules need not be ordered.

#### MISCELLANEOUS VOWEL ALTERNATIONS

Let us examine the miscellaneous vowel alternations in the following verbs:

	Present	Future	Stem	Imperative/Passive	
(13a)	<i>s-um-eelaq</i>	<i>mu-pu-seelaq</i>	<i>seelaq</i>	<i>suliq-i</i>	to kill
	<i>m-enaq</i>	<i>mu-m-eenaq</i>	<i>enaq</i>	<i>niq-i, niq-an</i>	to exist
(13b)	<i>mege</i>	<i>mu-bege</i>	<i>bege</i>	<i>biq-i, biq-un, biq-an</i>	to give
(13c)	<i>matís</i>	<i>mu-patís</i>	<i>patís</i>	<i>putas-i</i>	to write
(13d)	<i>m-adís</i>	<i>ma-adís</i>	<i>adís</i>	<i>des-i</i>	to bring

In (13a) vowel *i* appears in the penultimate stressed syllable, whereas vowel *a* appears in the unstressed final syllable. The same vowel alternation is also attested in the noun forms /seediq/ ~ /seedaq/ 'person' among different Sediq dialects. In both (13a) and (13b) the sequence /iq/ in the suffixed forms appears to be one of the male suffix forms that became attached to the noun or verb stems, with or without replacing part of the stem, in pre-Atayalic (see Li 1983). Most modern Atayal and Sediq speakers are unaware of the fact that /iq/ in these forms is a suffix. Another suffix can be attached to it, as in (13b). It is not clear, synchronically, how *i* is realised as *a* in the two verbs in (13a) and in the noun form /seedaq/.

The forms in (13c) show that the vowel changes in the opposite direction: *a* ---> *i*. The comparative evidence shows that the original vowel should be *a* rather than *i*.<sup>13</sup>

<sup>12</sup>According to Hsiu-fang Yang's (pers. comm.) transcriptions, the imperative forms for some of these verbs are: *duyag-i* 'Help!', *puyas-i* 'Sing!', *huyeg-i* 'Stand up!'. Apparently it varies from speaker to speaker whether the rule of vowel reduction (V ---> u/CV) or the rule of vowel assimilation (V ---> i/y/V) applies to some of these verbs.

<sup>13</sup>Cf. Atayal /matas/, Sediq (Toda dialect) /matas/, Rukai /wa-pacas/, Bunun /ma-patas/ 'to write', Saisiyat /pataS/ 'to tattoo', Thao /pataS-an/ 'book'.

Similarly, for the vowel alternation *e* ~ *i* as manifested in /des-i/ ~ /adis/ in (13d) the comparative evidence also shows that the original vowel should be *a* rather than *e* or *i*.<sup>14</sup>

All the irregular verbs in (13) present problems for a neat synchronic analysis. Again our diachronic evidence indicates that these verbs most likely had irregular verb inflections at an early stage.

#### BIBLIOGRAPHY

- LI, Paul Jen-kuei, 1977, Morphophonemic alternations in Formosan languages. *Bulletin of the Institute of History and Philology, Academia Sinica (BIHP)* 48/3:375-413.
- 1981, Reconstruction of proto-Atayalic phonology. *BIHP* 52/2:235-301.
- 1982, Linguistic variations of different age groups in the Atayalic dialects. *Tsing Hua Journal of Chinese Studies, New Series* 14/1-2:167-191.
- 1983, Types of lexical derivation in men's speech in Mayrinax. *BIHP* 54/3:1-18.
- YANG, Hsiu-fang, 1976, The phonological structure of the Paran dialect of Sediq (in Chinese). *BIHP* 47/4:611-706.

<sup>14</sup>Cf. Atayal /maras/, Sediq (other dialects) /madas/, Bunun /madas/, Saisiyat /maraS/, Kavalan /mazas/ 'to bring'.



limited. For the vowel alternation  $\epsilon \rightarrow \text{schwa}$  in (13) the computer program also shows that the original vowel should be a schwa in (13). All the irregular verbs in (13) present problems for a real synchronic analysis. The phonetic evidence indicates that these verbs must have had very different vowel qualities at an early stage.

#### REFERENCES

1. A. N. S. (1971). *Phonological analysis of the English lexicon*. London: Cambridge University Press.
2. The *Cambridge English Corpus* (1987-1992). Cambridge: Cambridge University Press.
3. The *Cambridge English Corpus* (1993-1998). Cambridge: Cambridge University Press.
4. The *Cambridge English Corpus* (1999-2004). Cambridge: Cambridge University Press.
5. The *Cambridge English Corpus* (2005-2010). Cambridge: Cambridge University Press.

## RECIPROCALLS AND DEPATIENTIVES IN TO'ABA'ITA

FRANTISEK LICHTENBERK

### 1. INTRODUCTION<sup>1</sup>

It is well known that there is an affinity between reflexives and reciprocals: in many languages the constructions that serve these two functions are historically related and may even be identical. Of the two functions, it is the reflexive one that has received considerably more attention in cross-linguistic, typological studies. When reciprocals are discussed in such works, it is usually in the context of reflexives (see, for example, Geniušienė 1987 and Kemmer 1988 and references therein). Reflexives are usually taken as primary, and reciprocals as secondary; the latter develop from the former, not vice versa. There is another sense in which the reflexive function is sometimes considered primary over the reciprocal function: reflexives are or are assumed to be the source of considerably more functional extensions than are reciprocals.

One of the purposes of the present study is to point out another aspect of the affinity between reflexives and reciprocals, and in so doing contradict the view that certain functions necessarily derive from the reflexive rather than the reciprocal function.<sup>2</sup> I will pay particular attention to the way in which the functional extension under discussion is motivated by the source, in this case the reciprocal, function.

The other main purpose of this study is an investigation of the nature of the function of the construction in question: is its function primarily syntactic, or is it primarily semantic/pragmatic, the syntactic properties being merely a reflection of the semantics/pragmatics? I will argue for the latter view.

The data for this study come from To'aba'ita, an Austronesian language spoken on Malaita Island in the south-east Solomon Islands.<sup>3</sup>

---

<sup>1</sup>It is with great pleasure that I dedicate this paper to George Grace. It was George, who (with Andrew Pawley) introduced me to Oceanic linguistics by inviting me (then a beginning doctoral student) to do fieldwork in Papua New Guinea. George was also my dissertation supervisor, and I could not have wished for a better one. His work both in Oceanic and in general linguistics has been a great inspiration to me.

I am grateful to Robert Blust for comments on an earlier version of this paper, in particular for those concerning the Proto Oceanic reciprocal prefix(es).

<sup>2</sup>It is, of course, possible that those functions have neither a reflexive nor a reciprocal but some other source.

<sup>3</sup>I am grateful to Lawrence Foanaota and Ruel Riianoa, my principal To'aba'ita consultants, for the data.

## 2. RECIPROCAL

To'aba'ita has several strategies to encode reciprocal situations. One of these consists in adding the prefix *kwai-* and, under some conditions, the suffix *-i* to the verb. The exact conditions governing the use of the suffix are not relevant here; suffice it to say that the suffix is not used if its presence would result in a new stress pattern in the verb. I will refer to this construction as the '*kwai* construction'; both the prefix and the suffix will be glossed 'R'.<sup>4</sup> The *kwai* construction is exemplified in (1):

- (1) *Roo wela kero kwai-kwa'e-i.*  
 two child they.DU R-hit-R  
 The two children hit each other.

Clauses that contain the *kwai* construction are syntactically intransitive: there is no direct object. Importantly for our present purposes, the *kwai* construction has other functions besides encoding reciprocal situations; one of these will be discussed in detail later.

Another way to encode reciprocal situations is by means of an independent personal pronoun in object position, coreferential with the subject. This 'simple pronominal construction' is exemplified in (2):

- (2) *Keero'a keko thathami keero'a...*  
 they.DU they.DU like them.DU  
 They liked each other...

The same construction is also used to encode reflexive situations; thus, out of context, (2) could also mean 'They liked themselves'. Furthermore, in (2) the object pronoun need not be bound by the subject: 'They<sub>i</sub> liked them<sub>j</sub> ...'.

Another alternative is to use an independent personal pronoun in object position as above, followed by the reciprocal marker *kwailiu*. *Kwailiu* consists of the prefix *kwai-* and *liu*, which otherwise means 'walk, take a walk, walk about'. The pronoun-plus-*kwailiu* construction is exemplified in (3):

- (3) *Roo wela kera fa'ama'u keero'a kwailiu.*  
 two child they frighten them.DU RECIP  
 The two children frightened each other.

*Kwailiu* has some other functions besides marking reciprocity. It is used with the meanings of 'back and forth' (see (4)), 'all over the place' (5), and 'respective(ly)' (6):

- (4) *'Osi liliu kwailiu.*  
 you.SG.NEG walk.RED back and forth  
 Stop walking back and forth! / Stop hanging around!
- (5) *'Osi 'uu kwailiu 'ana tafu 'ena.*  
 you.SG.NEG throw all.over.the place with rubbish that  
 Stop throwing the rubbish all over the place!

<sup>4</sup>The other abbreviations used in glossing the examples are: ART – article, ATTR – attributive, CLASS – classifier, CONTR – contrastive, DU – dual, I – inclusive, INT – intensifier, NEG – negative, NOM – nominaliser, PL – plural, RECIP – reciprocal, RED – reduplication, REL – relative-clause marker, SG – singular. Note that the plural forms of the subject markers are often used instead of the expected dual ones. The subject markers also encode tense or sequentiality; these are not indicated here.

- (6) *Kera oli kwailiu.*  
 they return respectively  
 They went back to their places (each person going back to his/her own place).

There is one more way in which reciprocal situations can be encoded, and that is by means of separate clauses:

- (7) *Thaari 'eri ka thathami-a tha wela 'eri ma wela 'eri*  
 girl that she like-him ART child that and child that  
*mena ka thathami-a la'u bo'o thaari 'eri.*  
 CONTR he like-her also INT girl that  
 The girl liked the boy (lit. child), and the boy, too, liked the girl.

Sentence (7) is taken from a text. The situation encoded in (7) is encoded twice more in subsequent sentences, where two of the other reciprocal strategies previously demonstrated are employed:

- (8)=(2) *Keero'a keko thathami keero'a...*  
 they.DU they.DU like them.DU  
 They liked each other...
- (9) *Keko thathami keero'a kwailiu.*  
 they.DU like them.DU RECIP  
 They liked each other.

In the discussion that follows, I will disregard encoding of reciprocal situations by means of separate clauses and will concentrate on the other three strategies.

The three monoclausal reciprocal constructions are different from each other in terms of their weight, where 'weight' refers both to the number of elements in the construction and to the lexical status of the elements. The more elements in a construction, the heavier the construction; and lexically independent elements are heavier than affixes. (The concept of 'weight' subsumes what Haiman 1983 refers to as 'linguistic separateness', in this case the distinction between bound morphemes and separate words: see also Kemmer 1988 for the notion of 'weight'.) Thus the *kwai* construction is the lightest: although it uses two elements, the two elements form one ambifix, and furthermore one of them (the suffix *-i*) is not always present. The construction that uses pronouns in object position and the reciprocal marker *kwailiu* is the heaviest: it uses two elements, both of which are separate lexemes. The construction with pronouns in object position but without *kwailiu* is of medium weight: it uses only one element, which, unlike *kwai...(-i)*, is a separate lexeme.

As examples (8) and (9) demonstrate, one and the same reciprocal situation can be encoded in different ways. In (10a-c) below, all three strategies are used to encode the same basic kind of situation (two participants standing in the same relation to each other):<sup>5</sup>

- (10) a. *Roo wela kera 'ui keero'a kwailiu 'ana fau.*  
 two child they pelt them.DU RECIP with stone  
 b. *Roo wela kera 'ui keero'a 'ana fau.*

<sup>5</sup>For the notion of reciprocal (and other kinds of) situations as consisting of relations see Lichtenberk (1985).

c. *Roo wela kera kwai-'ui-i 'ana fau.*

R-pelt-R

The two children threw stones at each other.

To say that all three strategies may be used to encode one and the same situation or the same basic kind of situation is not to say that they are fully synonymous and freely interchangeable; in fact they are not. The pronoun-plus-*kwailiu* construction has several uses. One of its functions is to make it explicit that the relations that make up a reciprocal situation are sequential rather than simultaneous (see examples (11) and (13) below). (Recall that *kwailiu* is also used to mean 'back and forth', which meaning is highly consistent with sequentiality of the relations in a reciprocal situation.) This sequential-marking function is applicable in (10a) above, but not in (9). The pronoun-plus-*kwailiu* construction is also used to emphasise the reciprocity of the relations in a situation; thus in (9) it emphasises that the liking was mutual. Notice that in the text, sentence (9) is the last in a sequence of three encodings of the situation. The multiple mentions, the use of *la'u* 'also' and the intensifier *bo'o* in the first mention (example (7)), and the use of *kwailiu* in the last mention (example (9)) all serve to emphasise the reciprocal nature of the situation. The pronoun-plus-*kwailiu* construction also serves to disambiguate reciprocals from reflexives and, in the case of third-person participants, from disjoint reference; as pointed out above, out of context, sentence (2) is three-way ambiguous. Finally, the construction is also used in cases where the *kwai* construction is not available (see example (18) below).

The *kwai* construction is normally used when the relations that make up a reciprocal situation are simultaneous or when their temporal configuration is irrelevant. This construction presents reciprocal situations as undifferentiated wholes, de-emphasising the existence of two (or more) separate relations within them and the separateness of the core participants. The core participants are conceptualised as members of the same set; they are encoded only once, in subject position (the construction is syntactically intransitive). The following examples serve to contrast the pronoun-plus-*kwailiu* and the *kwai* constructions. In (11), the preferred interpretation is sequential: first A helped B, and then B helped A:

- (11) *Kera 'adomi keero'a kwailiu.*  
 they help them.DU RECIP  
 They helped each other.

On the other hand, the normal interpretation of (12) is that the activities are more-or-less simultaneous, overlapping; in fact the sentence encodes a joint activity:

- (12) *Kulu kwai-'adomi 'ana na'are-laa.*  
 we.I R-help with cook-NOM  
 Let's do the cooking together. (Let's help each other with the cooking.)

In (12) it is not the case that both/all the persons will necessarily cook the same food at the same time and for the same period of time, rather that they will cooperate in preparing a meal: one may cook the rice, another may make the tea, etc. In fact, *kwai-'adomi* is normally used to mean 'cooperate', that is, 'help each other at (more or less) the same time'.

Compare also the next two examples. In the first one, the preferred interpretation is sequential:

- (13) *Roo wane kero thaungi keero'a kwailiu.*  
 two man they.DU hit them.DU RECIP  
 The two men hit each other. (First A hit B, then B hit A.)

On the other hand, in (14) the likely sequentiality of the individual events of punching is de-emphasised:

- (14) *Roo wane kero kwai-kumu-i.*  
 two man they.DU R-punch-R  
 The two men fought, exchanged blows.

Although, in principle at least, it is possible for the two men to have (always) punched each other simultaneously, a more natural interpretation of (14) is that there was a series of blows from both sides, the blows from the two directions interspersed with each other. The precise temporal configuration of the events is irrelevant; the situation is presented as an undifferentiated whole.

The third construction, with a personal pronoun but without *kwailiu*, lies in its use between the other two. Like the pronoun-plus-*kwailiu* strategy, the simple pronominal construction foregrounds the fact that there are two (or more) participants that play pairs of roles that stand in a mirror-image relation to each other, such as Agent and Patient. Second, like the pronoun-plus-*kwailiu* construction, the simple pronominal construction is used where the *kwai* construction is not available (see examples (17) and (19) below). On the other hand, like the *kwai* construction the simple pronominal construction is not used where sequentiality of the relations within a reciprocal situation is to be expressed or highlighted. All other things being equal (barring severe ambiguity and lexically-determined unavailability), it is the simple pronominal construction or the *kwai* construction that is used to encode reciprocal situations in which the relations are necessarily reciprocal, as in (15) and (16), taken from a text and referring to the same situation:

- (15) *...keki toda keero'a...*  
 they meet them.DU  
 ...they will meet ...
- (16) *...keki kwai-toda-i...*  
 they R -meet-R  
 ...they will meet ...

There are two other factors affecting the choice of a reciprocal construction, both of them having to do with the unavailability of the *kwai* construction. The *kwai* construction cannot be used when the source (i.e. non-reciprocal) verb requires an oblique rather than a direct object. In that case, one of the two pronominal strategies is required, with the pronoun as the object of the preposition:

- (17) *Roo kini kera mui uri keero'a.*  
 two woman they smile at them.DU  
 The two women smiled at each other.

Finally, there are also lexical restrictions on the use of the *kwai* strategy. This construction is not available with all the verbs that would otherwise be eligible. For example:

- (18) *Kera {lukata'i keero'a kwailiu/\*kwai-lukata'i.}*  
 they leave. for. good them.DU RECIP R-leave.for.good  
 They left each other for good. (Used to mean, they got divorced.)



- (19) *Roo wane kero*      (*labata'i keero'a/\*kwai-labata'i.*)  
 two man they.DU harm      them.DU R-harm  
 The two men harm each other.

The forms *kwai-lukata'i* and *kwai-labata'i* are possible, but they are not normally used to encode reciprocal situations. Other uses of the *kwai* construction, one of them in particular, will be discussed in section 3.

In a cross-linguistic investigation of the encoding of reciprocal (and other types of) situations, Kemmer (1988) introduces the notion of 'relative elaboration of events' that make up reciprocal (and other) situations. The degree of relative elaboration of events is a consequence of two factors: what Kemmer calls 'the distinguishability of participants' and 'the distinguishability of events'. The higher the distinguishability of the participants and/or the events, the higher the degree of elaboration. The notion of 'the distinguishability of participants' is related to what Langacker refers to as '(non-) distinctness of arguments' (Langacker and Munro 1975; Langacker 1976; see also section 4 below) and to what Hopper and Thompson (1980) refer to (following Timberlake 1975) as 'individuation of participants'. (For an application of the notion of 'individuation of participants' to reciprocal and some other types of situations see Lichtenberk 1985.) Haiman (1983) speaks of 'conceptual independence', which subsumes both distinguishability of participants and distinguishability of events. I will use the following terms to refer to these two concepts: 'the (degree of) distinctness of participants' and 'the (degree of) distinctness of relations'. Thus in a reciprocal situation, such as that encoded in 'John and Bill hit each other', the participants are less distinct from each other than they are in a simple transitive situation, such as that encoded in 'John hit Bill'. In the transitive situation, the two participants play different roles (Agent and Patient, respectively), whereas in the reciprocal situation they play identical pairs of roles (Agent-Patient). The notion of 'degree of distinctness of participants' is applicable even within reciprocal situations themselves. Arguably, the two participants in 'The two pieces of paper stuck together' (which in some languages would be encoded by means of a reciprocal construction) are less distinct from each other than the two participants in 'John and Bill hit each other'. As far as the degree of distinctness of relations is concerned, two like relations, such as those in 'John and Bill hit each other', are less distinct from each other than two unlike relations, such as those in 'John kicked Bill, and (then) Bill punched John'. In fact, it is because the relations in a reciprocal situation can be considered to be identical (apart from their directionality) that they can be encoded jointly by means of a single verbal form. The notion of 'degree of distinctness of relations' also is applicable within reciprocals. The relations that make up a reciprocal situation are less distinct from each other if they are simultaneous than if they were sequential. Furthermore, even relations that are not simultaneous can sometimes be conceptualised as non-distinct, with the overall situation treated as an undifferentiated whole. For example, in a situation describable by 'They conversed' (which in some languages requires a reciprocal construction) the individual events are more likely to be sequential (possibly with some overlap) rather than fully simultaneous; nevertheless, the overall situation is conceptualised as unary. The temporal configuration of the individual events is of no consequence (see also example (14) above). The notions of 'degree of distinctness of participants' and 'degree of distinctness of relations' will be relevant in the discussion of the functional extension of one of the To'aba'ita reciprocal constructions in section 4.

Together, the degree of distinctness of participants and the degree of distinctness of relations determine what, following Kemmer (1988), I will refer to as 'the (degree of) elaboration of

situations'.<sup>6</sup> By 'the degree of elaboration of situations' I mean the degree to which situations are both conceptualised and encoded as being comprised of distinct elements, that is, the participants and the relations among the participants. Any situation can be conceptualised and encoded in more than one way, paying more or less attention to detail, such as the internal structure of the situation. The degree of elaboration of a reciprocal situation is a direct consequence of the degree of distinctness of the relations comprising the situation and the degree of distinctness of the participants involved: the higher the degree of distinctness of the participants and/or the relations, the higher the degree of elaboration of the overall situation.

Following the work of Haiman (1983), Kemmer argues that if a language has more than one construction to encode certain kinds of situations (reciprocals among them) where constructions are of unequal weight, the uses of the constructions are motivated iconically, and she makes the following generalisation: "the extent of surface marking [i.e. the weight – FL] mirrors the conceptual dimension of degree of elaboration of events" (Kemmer 1988:165). With respect to reciprocals, Kemmer (pp.147-148) makes the following claim: in languages with light and heavy reciprocal constructions, "light reciprocal marking is strongly associated with simultaneity, while heavy reciprocal markers are temporally indifferent".

The To'aba'ita reciprocals fit the principle of iconic motivation rather well although there are some complications. All other things being equal, the lightest strategy (the *kwai* construction) is used for low degrees of elaboration: the relations as well as the participants are of a low degree of distinctness (the relations are simultaneous or their temporal configuration is irrelevant to the situation as a whole, and the core participants are treated as members of the same set). All other things being equal, the heaviest strategy (a pronoun in object position plus *kwailiu*) is used for high degrees of elaboration: the relations as well as the participants are distinct (the relations are sequential, and/or the reciprocity of the relations, that is, the existence of two relations, is made prominent, and so is the existence of the two participants each playing a pair of roles vis-a-vis each other). The medium-heavy strategy (an object pronoun alone, without *kwailiu*) is used for intermediate degrees of elaboration: the relations are not distinct but the participants are.

According to Kemmer, light reciprocal marking is associated with simultaneity while heavy reciprocal marking is temporally indifferent. However, in To'aba'ita, it is the two lightest strategies that are unmarked in terms of the temporal configurations of the constituent relations, while the heaviest strategy is marked, being associated with sequentiality. (The association of the heaviest strategy with sequentiality is not absolute because the construction has other functions: emphasis, disambiguation, and cases where the *kwai* construction is not available.) It remains to be seen whether Kemmer's strong generalisation can be upheld with To'aba'ita as an exception or whether the generalisation needs to be weakened to a negative one: no language with more than one reciprocal strategy of unequal weight consistently uses the heavier construction to encode simultaneity and the lighter construction to encode sequentiality.

The pronoun-plus-*kwailiu* strategy is more recent than the *kwai* strategy: it is found only in relatively close relatives of To'aba'ita, whereas the *kwai* construction is reconstructible at least as far back as Proto Oceanic (see section 4). It is conceivable that as the *kwai* construction began to acquire other functions (see section 3), a new reciprocal-marking strategy developed (a process called 'layering' by Hopper 1988). Although relevant historical evidence is lacking, it is likely that the

---

<sup>6</sup>Kemmer uses the term 'relative elaboration of events', but it is the elaboration of whole situations rather than that of the events internal to situations that is relevant here.

reciprocal function of *kwailiu* is an extension from the meaning 'back and forth' (see example (4) above), in which case the initial association of *kwailiu* with sequentiality would have been only natural.

### 3. DEPATIENTIVES

Besides encoding reciprocal situations, the *kwai* construction is used with other functions. One of these is the formation of what, following Kemmer (1988), one may refer to as the 'middle voice'. The middle-voice function is or at least appears to be lexically rather restricted, and at present no generalisations are possible. Following are a few examples of middle-voice verbs: *kwai-thathai* 'be ready (to do something)', cf. *thathai* 'make something ready'; *kwai-lukata'i* 'be close to death', cf. *lukata'i* 'leave somebody for good'; *kwai-kulufa'i* 'be infirm, walk with difficulty (of old people)', also 'be overcast (of weather)', cf. *kulufa'i* 'hang (vtr.)'; *kwai-liu* 'walk all over the place', also used as a reciprocal marker, and to mean 'back and forth', '(do something) all over the place', and 'respective(ly)', cf. *liu* 'walk, take a walk, walk about'. Kemmer suggests that cases where reciprocals give rise to the middle are not common (most frequently it is reflexives that give rise to the middle); To'aba'ita appears to be one such language.

The *kwai* construction has another non-reciprocal function, and it is this function I will focus on in the remainder of this paper. I will refer to this use of the *kwai* construction as 'depatientive'. In the depatientive function, the construction takes as its input a transitive verb, and the output is a syntactically intransitive verb. The subject of the derived verb corresponds to the subject of the source verb, but there is no direct object corresponding to the direct object of the source verb. For example:

- (20) *Oomea 'eri 'e kwai-fa'ama'u-i 'asiana'a.*  
 enemy that he R-frighten-R very  
 The enemy is very frightening.

Compare:

- (21) *'Osi faafa'ama'u-a wela 'ena.*  
 you.SG.NEG frighten.RED-him child that  
 Stop frightening the child!

(*Fa'ama'u* 'frighten' consists of *ma'u* 'be frightened' and the causative prefix *fa'a-* (reduplicated in (21).)

*Kwai-fa'ama'u-i* is used depatientively, but not reciprocally:

- (22) *Roo wela kera {fa'ama'u keero'a kwailiu/\*kwai-fa'ama'u-i.}*  
 two child they frighten them.DU RECIP R-frighten-R  
 The two children frighten each other.

As mentioned in section 2, there is a form *kwai-labata'i*, which is not normally used with a reciprocal function; instead, it is used depatientively:

- (23) *Roo wane kero kwai-labata'i.*  
 two man they.DU R-harm  
 The two men harm (people) (e.g. by stealing from them or by performing harmful magic on them).  
 Not: \*The two men harm each other. (See example (19).)

Also:

- (24) *si doo ki sui bana n-e kwai-labata'i.*  
 CLASS thing PL all REL-it R-harm  
 all the harmful things, all the things that harm

Compare:

- (25) *'Aburu na'i labata'i koro 'asiana'a.*  
 giant this harm us.I.DU very  
 The giant harms us greatly.

Further examples of the depatientive use of the *kwai* construction:

- (26) *Wane 'e kwai-ilamata'i.*  
 man he R-hate  
 The man hates people/everybody/everything. / The man is a hater.

Compare:

- (27) *Wane na'i 'e ilamata'i nau.*  
 man this he hate me  
 This man hates me.

- (28) *wane ni kwai-fa'amaruki*  
 man ATTR R-save/revive  
 Saviour ('man' who saves people/lives)

Compare:

- (29) *...ka fa'amaruki nia bia roo wela nia ki.*  
 it save/revive her and two child her PL  
 ...and it saved her and her two children.

Notice that while in the reciprocal construction the subject cannot be (semantically) singular (it takes at least two participants to act on each other), in the depatientive construction there is no prohibition against singular subjects.

Clauses with the depatientive construction are syntactically intransitive; there is no direct object. At the same time, such clauses are transitive semantically. Although no patient/direct object is expressed, there is always one implied. The implied patient is general, non-specific; the event encoded by the verb is directed not at a specific participant but at any and all of a certain kind.

One might want to suggest that in the depatientive function the *kwai* construction functions as an intransitiver. It is true that the construction is used with transitive sources and that the output is intransitive, but this can hardly be its (real) function. Grammars exist to enable their users to encode meanings, not to instantiate formal properties and relations. To a significant degree, the properties of grammars are motivated by semantic and/or pragmatic factors. With certain (originally) transitive verbs, if the patient is of low salience, it is backgrounded, made non-prominent through not being encoded at all. The formally intransitive nature of clauses with the depatientive construction is a consequence of the semantics/pragmatics of the situations encoded.

## 4. THE RECIPROCAL-DEPATIENTIVE CONNECTION

What I call the 'depatientive' construction is far from rare in the languages of the world. For example, Geniušienė (1987) mentions parallel constructions as existing in a number of geographically as well as genetically diverse languages. She terms such verbs 'absolute reflexive verbs'. Geniušienė's study takes as its point of departure (true) reflexives and then investigates the extension of (originally) reflexive markers into other functional domains, reciprocal and absolute-reflexive/depatientive among them. In several places in her study, Geniušienė mentions the coexistence of the reciprocal and the depatientive functions, both marked in the same way. Where the same construction is (or used to be) also used to mark reflexive situations, Geniušienė assumes that the reciprocal and the depatientive functions are independent extensions (directly or indirectly) from the reflexive function. She does mention that in (some) Turkic languages one and the same marker is used to encode reciprocal and depatientive situations, plus some others, but not reflexive situations. She assumes that the depatientive function is an extension from the reciprocal function.

The situation in To'aba'ita is comparable to that in the Turkic languages. The *kwai* construction has a reciprocal and a depatientive function but not a reflexive function, and, as far as one can tell, it did not have a reflexive function when the depatientive function developed (see below). Since it is possible for the reciprocal and the depatientive functions to develop one from the other in the Turkic languages and in To'aba'ita (and its close relatives), it is not impossible for similar developments to have occurred in languages where the same construction also has (or used to have) a reflexive function.

Geniušienė (1987:138) suggests that the reciprocal-depatientive polysemy is "determined by the fact that the meaning and diathesis of ... [the source verbs - FL] ... allows for a change in two ways" from an earlier reflexive source, but she does not say what the relevant factors might be. Without wishing to deny the relevance of verb meanings to functional extensions of grammatical elements, I suggest that what is more important is the similarity between the internal structures of the two types of situation encoded by the construction, as conceptualised by language users. If two types of situation are similar in some relevant respects, it is possible for a construction originally used to encode only one situation type to be extended to encode the other as well.

As far as one can tell from the comparative evidence, the reciprocal-depatientive polysemy of the *kwai* construction is the result of an extension from the reciprocal to the depatientive function. Pawley (1973) has reconstructed a reciprocal prefix *\*paRi-* for Proto Oceanic, and Blust (forthcoming and pers. comm.) has reconstructed *\*paRi-* for Proto Austronesian as a 'prefix of reciprocal or collective action'. Pawley suggests that besides *\*paRi-* a doublet form *\*pai-* is to be reconstructed for Proto Oceanic (see also Ross 1988); on the other hand, Blust argues that the witnesses suggesting Proto Oceanic *\*pai-* are in fact reflexes of *\*paRi-* that have undergone irregular loss of the *R* (grammatical morphemes being prone to sporadic reduction). Whatever the case may be, there is evidence that Proto Oceanic reciprocal marking consisted not only of *\*paRi-/pai-* but, at least in some cases, also of a suffix *\*-i* (Pawley 1973). (For example, Manam has a suffix *-i* used together with the reciprocal prefix *e-* (a reflex of *\*paRi-/pai-*) to encode those reciprocal situations where there is 'transfer of an activity' between the participants (Lichtenberk 1983:211): *e-boabu-i* '(briefly) embrace each other', but *e-boabu* 'clasp each other'.) In the To'aba'ita *kwai* construction, the suffix *-i* reflects the Proto Oceanic suffix *\*-i*. However, the etymology of the prefix *kwai-* is less clear. The prefix exhibits some formal irregularities vis-a-vis the Proto Oceanic form(s): rather than *kwai-*, the expected form is *\*\*fali-* (from *\*paRi-*) or *\*\*fai-* (from *\*pai-*). The same irregular forms are also



found in close relatives of To'aba'ita, for example, Lau, Kwaio and Kwara'ae.<sup>7</sup> On the other hand, some slightly more distant relatives of these languages have reflexes both of the \*paRi- type and of the pai- type, for example, Arosi *hari-* and *hai-* (Fox 1978), and Nggela *vari-*, *vei-* and even *vaivari-* (Fox 1955).<sup>8</sup> Whatever the exact historical development of To'aba'ita *kwai-* may have been, given the fact that Proto Oceanic had a reciprocal construction of the form \*paRi-/pai-...-i and that the To'aba'ita reciprocal construction has the (basic) form *kwai-...-i* one can be fairly certain that the To'aba'ita construction does continue the Proto Oceanic construction(s).

There is abundant evidence for positing a reciprocal-marking function for Proto Oceanic \*paRi-/pai-...-i but no evidence for assuming a reflexive-marking function. Oceanic cognates of the *kwai* construction with the depatientive function are found in languages closely related to To'aba'ita, such as Kwara'ae (Deck 1934), Kwaio (Keesing 1985) and most likely Lau (Fox 1974), and also in Fijian, a member of a different primary subgroup of Oceanic (Schütz 1985). The reciprocal function of the reflexes of the Proto Oceanic construction(s) is much more widely distributed in the family than the depatientive function; in fact, in every language that has a reflex the construction has a reciprocal function and perhaps some others, but not necessarily a depatientive function. The evidence then points to the following conclusions: the depatientive function is a later, post-Proto-Oceanic development; its presence in different first-order subgroups of Oceanic is due to independent developments; and at the time when the reciprocal-to-depatientive extension took place, the construction(s) did not have a reflexive function. I am not suggesting that the extension necessarily took place in To'aba'ita. The same extension is also found in close relatives of To'aba'ita; interestingly, those languages also use cognates of To'aba'ita *kwailiu* with a reciprocal function. Even though the possibility of independent developments cannot be discounted, it is quite likely that the extension of the *kwai* construction into the depatientive function as well as the extension of *kwailiu* (from 'back and forth') into the reciprocal function took place at a pre-To'aba'ita stage. Here I am interested not in determining when the depatientive extension took place, but in the nature of the extension, in the factors that motivate it, which, one can assume, would not have been appreciably different no matter when the extension took place. Finally, the main thrust of the argument developed in what follows (that the reciprocal-depatientive polysemy is due to similarity between the two types of situation) would not be in any serious sense invalidated if it should turn out that it was the depatientive function that was historically primary: if A is similar to B, then B is similar to A.

The basic factor that motivates the extension from the reciprocal to the depatientive function is the relatively low degree of elaboration of situations due to the low degrees of distinctness of both the relations and of the participants. Crucially, there are similarities between the two situation types with respect to the specifics in which they are of low elaboration. (The notion of 'relative elaboration of situations' is so general that there must presumably be fairly specific similarities between two situation types for a functional extension to take place. A priori it seems unlikely that a functional

<sup>7</sup>There is one further irregularity found in connection with the prefix *kwai-*. The prefix (in To'aba'ita and the other close relatives) has a doublet *wai-*, which is used in reciprocal (and certain other) kinship terms, for example, To'aba'ita *waithaulana* 'people of same sex married to siblings' (e.g. two men married to sisters). *kw* reflects earlier (Proto-Central-North-Malaitan) *w* (Lichtenberk 1988). For whatever reason, the prefix \**wai-* did not undergo the *w*-to-*kw* change in its kinship function.

<sup>8</sup>The multiple reflexes need not be (fully) synonymous. Thus, according to Fox (1978:184), Arosi "*hai* differs from *hari* in that the former usually shows mutual assisting or conflicting action and *hari* merely combined action, but there are exceptions in both cases".



extension would take place if the specific factors responsible for a low degree of elaboration in two situation types were completely different from each other.)

There are a number of factors relevant to the degree of distinctness of the relations in situations encoded by reciprocals and depatientives. Two relations that are of the same type (two events of kicking) are less distinct from each other than two different relations (one event of kicking and one of feeding). A relation occurring only once is more distinct, more prominent than any of a number of what can be considered to be the same relations occurring simultaneously or in (close) sequence. And, all other things being equal, relations occurring simultaneously or those whose exact temporal configuration vis-a-vis each other is irrelevant also are relatively non-distinct.

As far as participants are concerned, two participants are less distinct from each other if they play identical roles or sets of roles in the relations they are involved in than if their roles were different. Secondly, two participants are relatively non-distinct by virtue of suppression of their ultimate distinctness. This view of relative non-distinctness of participants is based on Langacker's notion of 'non-distinct arguments' (Langacker and Munro 1975; Langacker 1976). In Langacker's framework, if the referent of an underlying argument is not overtly specified, that argument is non-distinct from another, specified argument. One need not subscribe to Langacker's fairly abstract syntactic analysis to be able to adopt his insight concerning non-distinctness; secondly, in my approach, it is better to speak of relative (non-)distinctness of participants rather than arguments. Two participants one or both of which are of low salience in a given situation are less distinct from each other than they would be if both were highly salient.

In reciprocal situations encoded by the *kwai* construction, the participants are of a low degree of distinctness from each other: both (all) the participants play identical pairs roles; in fact, they are conceptualised as members of the same set (they are encoded only once). The relations are identical (apart from their directionality), and temporally they are non-distinct. The overall situation is conceptualised as an undifferentiated whole.

With depatientives, although the two core participants are involved in the relations in different roles, one of the participants is of low salience, which makes the two relatively non-distinct from each other. The relations that make up the situations encoded by the depatientive construction are general, habitual rather than specific; the construction expresses a characteristic relation of the subject participant toward any and all entities of a certain kind ('hater', 'saviour', 'harmer' etc.) Thus in this kind of situation complex, there are multiple relations, which by virtue of the low salience of one of the participants are relatively non-distinct from each other (all events of, for example, 'harming' are on a par, regardless of the exact identity of the participants they are directed at). The depatientive construction encodes general, habitual situations, where the temporal configuration of the internal relations is of no consequence. The relations are temporally non-distinct from each other; the whole situation complex is treated as an undifferentiated whole. It is these specific similarities between the two types of situation that underlie and motivate the rise of the reciprocal-depatientive polysemy.

In the depatientive construction, the focus is on the subject participant. The salience of the other participant is very low; in essence, that participant is part of the background. This is reflected in the structure of the clause: the non-salient participant is not encoded at all; the clause is syntactically intransitive. The depatientive construction is motivated iconically: lack of salience results in lack of overt encoding.

Of course, to say that a reciprocal-depatientive polysemy is natural because it is motivated is not to say that it is necessary. After all, there are many languages in which the two types of situation do not

receive the same, or historically related, marking. Nevertheless, as a growing number of cross-linguistic, typological studies have demonstrated, motivation is an important factor in the shape of the grammars of individual languages, and as such it is a crucial concept for our understanding of grammar. The intransitivising property of the *kwai* construction both in the reciprocal and in the depatientive functions is not an arbitrary formal property; it is a consequence of the use of the construction to encode certain types of situation.

## BIBLIOGRAPHY

- BLUST, Robert A., Forthcoming, *Austronesian comparative dictionary*.
- DECK, Norman C., 1934, *Grammar of the language spoken by the Kwará'ae people of Mala, British Solomon Islands*. *Journal of the Polynesian Society* Reprint 5. Wellington: The Polynesian Society.
- FOX, Charles E., 1955, *A dictionary of the Nggela language (Florida, British Solomon Islands)*. Auckland: Unity Press.
- 1974, *Lau dictionary*, with English index. *PL*, C-25.
- 1978, *Arosi dictionary* (Revised edn.) *PL*, C-57.
- GENIŠIENĖ, Emma, 1987, *The typology of reflexives*. *Empirical Approaches to Language Typology* 2. Berlin, New York and Amsterdam: Mouton de Gruyter.
- HAIMAN, John, 1983, Iconic and economic motivation. *Language* 59/4:781-819.
- HOPPER, Paul, 1988, Some heuristic principles of grammaticization. Paper presented at the Symposium on Grammaticalization, University of Oregon, May 12-15, 1988.
- HOPPER, Paul J. and Sandra A. THOMPSON, 1980, Transitivity in grammar and discourse. *Language* 56:251-299.
- KEESING, Roger M., 1985, *Kwaio grammar*. *PL*, B-88.
- KEMMER, Suzanne, E., 1988, *The middle voice: a typological and diachronic study*. PhD dissertation, Stanford University.
- LANGACKER, Ronald W., 1976, *Non-distinct arguments in Uto-Aztecan*. University of California Publications in Linguistics 82. Berkeley and Los Angeles: University of California Press.
- LANGACKER, Ronald W. and Pamela MUNRO, 1975, Passives and their meaning. *Language* 51:789-830.
- LICHTENBERG, Frantisek, 1983, *A grammar of Manam*. *Oceanic Linguistics* Special Publication 18. Honolulu: University of Hawaii Press.
- 1985, Multiple uses of reciprocal constructions. *Australian Journal of Linguistics* 5:19-41.
- 1988, The Cristobal-Malaitan subgroup of Southeast Solomonian. *Oceanic Linguistics* 27:24-62.
- PAWLEY, Andrew, 1973, Some problems in Proto-Oceanic grammar. *Oceanic Linguistics* 12:103-188.
- ROSS, M.D., 1988, *Proto Oceanic and the Austronesian languages of Western Melanesia*. *PL*, C-98.
- SCHÜTZ, Albert J., 1985, *The Fijian language*. Honolulu: University of Hawaii Press.
- TIMBERLAKE, Alan, 1975, Hierarchies in the genitive of negation. *Slavic and East European Journal* 19: 123-138.

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: 773-936-3200  
WWW.CHICAGO.LIBRARY.EDU

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: 773-936-3200  
WWW.CHICAGO.LIBRARY.EDU

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: 773-936-3200  
WWW.CHICAGO.LIBRARY.EDU

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: 773-936-3200  
WWW.CHICAGO.LIBRARY.EDU

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: 773-936-3200  
WWW.CHICAGO.LIBRARY.EDU

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: 773-936-3200  
WWW.CHICAGO.LIBRARY.EDU

## A CENTURY OF LINGUISTIC CHANGE IN ANEJOM

JOHN LYNCH

### 1. INTRODUCTION

It is a pleasure to offer this small contribution to a volume in honour of George Grace. His influence on the field of Austronesian linguistics has been immense, and has perhaps been felt most noticeably in those parts of Melanesia where the more 'aberrant' languages are spoken. Anejom, or Aneityumese, is one of those languages, and the data and issues discussed here are of a kind which George has addressed in other parts of Melanesia.<sup>1</sup>

In 1882, the Rev. John Inglis published his Anejom dictionary, which included a brief grammatical sketch. Anejom thus became one of the first languages in Oceania for which we have reasonable grammatical and lexical material. Very little work on the language was carried out after 1882, however; most subsequent publications which refer to it are of a comparative nature, and are based either on Inglis's published work or on the mission translations of religious materials.<sup>2</sup> I use the term *Old Anejom* to refer to this stage of the language.

Exactly one hundred years after the publication of Inglis's dictionary and grammar sketch, I prepared a grammar outline (Lynch 1982a) and a lexical file (Lynch 1982b), the latter based both on my own fieldwork and on two other relatively recent publications (Hewitt 1966; Tryon 1976), all describing what I refer to here as *Modern Anejom*. We are thus in a reasonable position to examine at least some of the changes which have taken place in one Oceanic language over the period of a century, to assess the nature of those changes, and to examine their possible causes.

Anejom is spoken by about five hundred people, most of whom live on the island of Aneityum, the most southerly inhabited island in the Republic of Vanuatu. Only one Melanesian language is spoken on the island, and even at the time of European contact this seems to have been the case, although at that time the degree of dialect differentiation appears to have been greater than it is today. (The population at that time was also much larger, but due to a combination of imported diseases,

---

<sup>1</sup> An earlier version of this paper was given as a seminar at the Australian National University. I am grateful for comments received there – especially from Tom Dutton and Francesca Merlan – and also for comments on this version received from Terry Crowley.

<sup>2</sup> Abbreviations used in this paper are: 1,2,3 – first, second, third persons, AOR – aorist, DU – dual, EXC – exclusive, FUT – future, INC – inclusive, NEG – negative, NON-SG – non-singular, OBJ – objective, PAST – past, PERF – perfective, PL – plural, SG – singular, SUBJ – subjunctive, SUBJ-MKR – subject-marker, TEMP – temporal, TR – transitive, TRL – trial.

natural disasters, and rabid nineteenth-century Presbyterianism, there has been a very severe decline in the total population of the island.)

This paper will concentrate on three areas in which changes have taken place: the lexicon, the numeral system, and the system of preverbal particles which mark subject, tense and mood. I begin, however, with a brief outline of Anejom phonology.

## 2. PHONOLOGY

The phonemes of Anejom (Hewitt 1966; Lynch 1982a) are given in Table 1. Symbols which may require some elucidation are: /b/ = [p<sup>w</sup>], [b<sup>w</sup>]; /j/ = [tʃ], [dʃ], [tʃ], [dʒ]; /d/ = [θ]; /c/ = [ç]; /m̃/ = [m<sup>w</sup>]; and /g/ = [ŋ].

TABLE 1: ANEJOM PHONEMES

Consonants		Stops	Fricatives	Nasals	Liquids	Semi-vowels
Velarised labials		<i>b</i>		<i>m̃</i>		<i>w</i>
Labials		<i>p</i>	<i>f, v</i>	<i>m</i>		
Dentals		<i>t</i>	<i>d, s</i>		<i>l</i>	
Alveolars				<i>n</i>	<i>r</i>	
Palatals		<i>j</i>		<i>ñ</i>		<i>y</i>
Velars		<i>k</i>	<i>c</i>	<i>g</i>		
Post-velars			<i>h</i>			
Vowels		Front	Central		Back	
High		<i>i</i>			<i>u</i>	
Mid		<i>e</i>			<i>o</i>	
Low			<i>a</i>			

It is difficult to assess absolutely accurately whether there have been any *significant* changes in the Anejom phonological system in the course of the last century. The major reason for saying this is that Inglis did not consistently represent various phonemic distinctions. Thus he did not distinguish the velarised and simple bilabials (i.e. /m̃/ from /m/, or /b/ from /p/), despite his use of the orthographic symbol *b*, presumably to represent (inconsistently) the voiced allophone of /p/. He also wrote as *i* the postvocalic high glide which predictably precedes any palatal consonant: thus he writes what is phonemic /esej/, phonetic [ʔeseitʃ] 'three', as *eseij*. While he did not use any special symbol for the palatal nasal /ñ/, it is clear that it was present in the language in his time, as his use of non-final *ny* and *ni* and final *ig* for phonemic /ñ/ testifies. My feeling is that these omissions from Inglis's sketch are just that – omissions rather than recent developments: that is, I believe he either did not hear or, if he heard it, did not orthographically represent the difference between /b/ and /p/, between /m̃/ and /m/, and between /ñ/ and combinations of vowel with /n/ or /g/.

In view of this, it is probably safe to say that there have been no major phonological changes in Anejom since 1882.

## 3. LEXICAL CHANGES

Inglis's dictionary contains between three and four thousand Anejom lexical items. Since his time, little lexical research has been carried out on the language: Hewitt (1966) published over twelve hundred items, Tryon (1976) lists around three hundred, and I collected in the vicinity of a thousand; many of the items in these three collections, obviously, are the same. These three lists have been amalgamated into an Anejom lexical file (Lynch 1982b), as yet unpublished. Thus there is a reasonable basis for lexical comparison, although in the absence of a large database of Modern Anejom lexical items it is difficult to come to any firm conclusions.

In this paper I will restrict my discussion to obvious borrowings from one or other of the major contact languages in Vanuatu – English, French or Bislama (the Vanuatu 'dialect' of Melanesian Pidgin). I will ignore what might be called 'necessary' borrowing, and focus on 'unnecessary' borrowing (see Clark 1982). By this I mean that I will not be concerned with borrowed words whose referents are totally new items in Anejom culture – words such as *nariko* 'bean', *sinema* 'cinema' and *taun* 'town'. Rather, I will be concerned with words which have been borrowed when their referents could be presumed to already exist, in some form or another, in traditional Anejom culture.

The first group of words I want to consider are borrowings which appear both in Inglis's and my own lists. These represent words borrowed from English or Bislama before 1882, and words for which one might have expected a 'native' Anejom equivalent. Only two words clearly fall into this category. The first is *lav* 'to laugh' (with an alternate *laav* (Inglis 1882)). Inglis also gives a form *ehehei* 'to laugh loudly' (probably onomatopoeitic), but *lav* appears even in 1882 as the basic verb 'to laugh'. The second is the form *naifi* 'knife'. Inglis records *nau itoga* 'knife' (lit. 'knife foreign'), but also *naife asvahteuc* 'shutting knife' (lit. 'knife bend'), probably referring to a pocket-knife or some other kind of knife with a folding blade. In Modern Anejom, both *nawutoga* and *naifi* are used; note also Modern Anejom *nau-ata* 'circumcision knife'.

TABLE 2: 'UNNECESSARY' BORROWINGS FROM BISLAMA AND/OR ENGLISH  
SINCE 1882

Old Anejom (Inglis 1882)	Modern Anejom (Lynch 1982b)
1. <i>nipjiuru</i> iron pot <i>nipjinitai</i> vessel, dish	<i>baila</i> pot [ <i>nipjin-nitai</i> bowl, dish]
2. <i>nahtaicai</i> gourd <i>nerero</i> calabash	<i>bañjeg</i> pumpkin [ <i>nerero</i> gourd]
3. <i>auaredo</i> play <i>didimai itai</i> play at games	<i>eplei</i> play (esp. European-style games) [ <i>ikru</i> play (traditional games)]
4. <i>adikjai, adikdikjai, edidikjai, alealeaujai</i> to dance <i>aurupu</i> clap hands	<i>taanes</i> to dance, <i>intaanes</i> a dance – both referring to European-style dances [ <i>autupu</i> to dance (a traditional dance)]
5. <i>naman</i> tongue (= ? <i>nama-n</i> his tongue)	<i>toog</i> tongue [ <i>nama-</i> tongue (archaic)]
6. <i>namasaki</i> leprosy	<i>nalepres</i> leprosy
7. <i>inceen nipan (napan?)</i> mast (lit. trunk bark)	<i>mas, inci-mas</i> mast



8. <i>nelcau</i> canoe	<i>kinou</i> canoe [ <i>nelcou</i> canoe, ship]
9. <i>nelcaupsu</i> a comb <i>ucji</i> to comb	<i>koom</i> a comb [ <i>inticjijinatimi</i> a comb, <i>ecjii</i> to comb]
10. <i>natohtan</i> insect like a large bee	<i>inhone</i> bee
11. <i>intihtag</i> house in village or district where people meet	<i>nakamal</i> /nakamal (from Bislama: a place where men gather to drink kava, and also a place where dances and other village festivities are held) [ <i>inteptag</i> dance-ground]
12. <i>nitai ahre</i> broom	<i>neprum</i> broom [ <i>intaarei</i> , <i>nitaaremne</i> broom]

The second group of words is given in Table 2. These are forms which appear to have been introduced since the publication of Inglis's dictionary; at least, they occur as 'normal' words in the modern language but are not listed by Inglis. The left column contains Old Anejom forms for which the Modern Anejom equivalents are given in the right column. In dealing with the Modern Anejom forms, the borrowing is given first in each case; where relevant, I have enclosed in square brackets either the continuation of the form cited by Inglis, and/or other Modern Anejom forms which are synonymous or nearly so and which are not borrowings from English or Bislama.

Some of these forms may be explained as borrowings which refer to a particular kind of item. The case of *naifi*, discussed above, is a good example, since the original Inglis data show that *nau* (or compounds with *nau*) probably referred to traditional knives made from bamboo (note Modern Anejom *nawunau* 'bamboo'), while *naifi* clearly first referred to a special kind of knife – either one which folded or one made from some kind of metal, but in any case a kind of knife which was imported.

Others in this category may well include *baila* 'pot' (specifically for boiling food?), *bañjeg* 'pumpkin' (as opposed to other gourd-like fruits), *eplei* 'to play European-style games', *taanes* 'to dance European-style dances', and *inhone* 'bee' (i.e. insect which produces honey).

The other forms are not so easily explicable. Certainly *lav* 'to laugh', *toog* 'tongue', *kinou* 'canoe', *nakamal* 'nakamal', *neprum* 'broom', and probably also *nalepres* 'leprosy', *mas* 'mast' and *koom* 'a comb', might well be expected to have had indigenous equivalents in Anejom prior to European contact. The reasons for borrowing in these cases remain obscure.

#### 4. CHANGES IN THE NUMERAL SYSTEM

Like a number of languages in Vanuatu and, indeed, in the rest of Oceania, Anejom has undergone considerable change in its numeral system. Numerals above 'five', which earlier consisted of quite lengthy compounds, have been replaced by borrowings from Bislama for reasons of convenience: for example, it is more 'convenient' to say *sikis* than *nikman celed et ethi* for 'six'. However, these are not the only changes which have taken place in the Anejom numeral system, as Table 3 shows.

TABLE 3: CHANGES IN THE ANEJOM NUMERALS SINCE 1882

	Inglis (1882)	Capell (n.d.)	Lynch (1982a)
one	<i>ethi</i>	<i>ethi</i>	<i>ithii</i>
two	<i>ero, ohwat</i>	<i>ero</i>	<i>erou</i>
three	<i>eseij</i>	<i>eseij</i>	<i>esej</i>
four	<i>emanowan</i>	<i>manohwan</i>	<i>mijman</i>
five	--	<i>nikman</i>	<i>meled</i>
six		<i>nikman celed et ethi</i>	[Bislama forms are now used]
seven		<i>nikman celed et ero</i>	
eight		<i>nikman celed et eseij</i>	
nine		<i>nikman lep ikman</i>	
ten		<i>nikman ero un reduon</i>	

We can, I believe, ignore as insignificant the spelling differences in the forms of the numerals 'one', 'two' and 'three'. We cannot ignore the form *ohwat*, given by Inglis as an alternative to *ero* 'two', but since he gives no further details, and since Capell does not mention it, we must just put it aside as a curious and (at the moment at least) unsolvable puzzle.

The forms for 'four' and 'five', however, are more interesting in that a systemic change seems to have taken place. Modern Anejom *mijman* 'four' seems to derive from earlier *nikman* 'five' and has entirely replaced the earlier form *emanowan* (Inglis) or *manohwan* (Capell); note that both Inglis and Capell give the form *nikman* 'hand, five' for what Hewitt, Tryon and I recorded as *nijman* 'hand', *mijman* 'four'. Similarly, Modern Anejom *meled* 'five' seems to derive from earlier *celed*, the linking particle for numerals above 'five'.

An almost identical systemic change has taken place in Ura, a moribund language of northern Erromango. There, earlier Ura *lemelu* 'four', *suorem* 'five' have been replaced by later Ura *suorem* 'four', *misikai* 'five'; cf. earlier Ura *misai* 'six' (Lynch 1983:150-151). It is interesting to speculate why two related but geographically relatively distant languages would have made the same unusual change, while other (geographically intervening) members of the Southern Vanuatu family have made no such change. The change itself, however, remains unexplained.

## 5. CHANGES IN THE VERBAL SYSTEM

By far the most interesting linguistic changes in Anejom are those which have taken place in a particular set of preverbal particles. A 'normal' Anejom transitive sentence consists of verb phrase + object NP + subject NP; a 'normal' intransitive sentence consists of verb phrase + subject NP. An Anejom verb phrase consists of a verb (with transitive or pronominal-objective suffixes where required), preceded by a number of sets of particles. For example (in Modern Anejom):

*Ek pu hag aňak.*  
 1SG.AOR FUT eat I  
 I will eat.

*Et itiyi fi hag aen.*  
 3SG.AOR NEG not.yet eat he  
 He hasn't eaten yet.

*Is mān lep ege-yin attaj.*  
 PAST PERF again hear-3SG.OBJ they.TRL  
 The three of them heard him again.

In this paper I will be examining the particles which mark subject, tense/aspect, and mood, and which appear as the first particle in any verb phrase. I will not be concerned with other preverbal particles which mark negativity, certain other aspects, or other adverbial functions.

For convenience of presentation, I distinguish four tense/aspect markers in Anejom:<sup>3</sup> aorist (marking present and recent past), future, past and subjunctive (used in certain kinds of conditional clauses). These forms, in many cases, combine with following aspect markers to encode other aspects. For example:

*Et apam aen.*  
 3SG.AOR come he  
 He is coming. / He just came.

*Et mān apam aen.*  
 3SG.AOR PERF come he  
 He has come.

*Et wut apam aen...*  
 3SG.AOR TEMP come he...  
 When he comes...

Anejom pronouns, in their free/focal, objective and possessive forms, distinguish four persons (first inclusive, first exclusive, second and third) and four numbers (singular, dual, trial and plural). In Old Anejom, these person and number distinctions were also reflected in the preverbal particles in which we are interested, as Table 4 shows.

TABLE 4: OLD ANEJOM PREVERBAL PARTICLES

	Singular	Dual	Trial <sup>4</sup>	Plural
Aorist				
1 INC		<i>intau</i>	<i>intaj</i>	<i>inta</i>
1 EXC	<i>ek</i>	<i>ecrau</i>	<i>ektaj</i>	<i>ecra</i>
2	<i>na</i>	<i>ekau</i>	<i>ahtaj</i>	<i>eka</i>
3	<i>et</i>	<i>erau</i>	<i>ehtaj</i>	<i>era</i>
Future				
1 INC		<i>intau pu</i>	<i>intaj pu</i>	<i>intu pu</i>
1 EXC	<i>ek pu</i>	<i>ecrau pu</i>	<i>ektaj pu</i>	<i>ecru pu</i>
2	<i>na pu</i>	<i>aka pu</i>	<i>ahtaj pu</i>	<i>aku pu</i>
3	<i>et pu</i>	<i>era pu</i>	<i>ehtaj pu</i>	<i>eru pu</i>

<sup>3</sup>The future tense/aspect marker is essentially the aorist tense/aspect + *pu*, with some changes in vowel quality in the aorist particles; but it is included here to show both these morphophonemic changes and to illustrate the kinds of historical changes that have taken place in Anejom. Note also that the terms 'aorist' and 'subjunctive' are not used by Inglis or Capell, but they do seem to me to be the best labels for these tense /aspect markers.

<sup>4</sup>Inglis (1882) showed no person distinctions in any of the trial preverbal particles; Capell (n.d.), however, gives person distinctions for aorist, future and past.

## Past

1INC		<i>intis</i>	<i>intijis</i> <sup>5</sup>	<i>intis/imjis</i> <sup>6</sup>
1EXC	<i>kis</i>	<i>ecrus</i>	<i>ektijis</i>	<i>ecris</i>
2	<i>as</i>	<i>akis</i>	<i>ahtijis</i>	<i>akis</i>
3	<i>is</i>	<i>erus</i>	<i>ehtijis</i>	<i>eris</i>

## Subjunctive

1INC		<i>tu</i>	<i>tiji</i>	<i>ti</i>
1EXC	<i>inki, ki</i>	<i>ecru</i>	<i>tiji</i>	<i>ecri</i>
2	<i>an</i>	<i>eru</i> <sup>7</sup>	<i>tiji</i>	<i>aki</i>
3	<i>inyi</i>	<i>eru</i>	<i>tiji</i>	<i>ari</i>

The data in Table 4 represent information given by both Inglis (1882) and Capell (n.d.). Capell says of the forms under discussion that they were "no doubt originally independent particles, but [are] now disguised by various degrees of compounding with elements of a pronominal nature" (n.d.:60). It is possible to propose an analysis of the Old Anejom particles which may reflect an even earlier stage in the language when the compounding referred to by Capell had not yet taken place. The following analysis is suggested by the data in Table 4 (with V symbolising a vowel which may well have been present, but whose quality cannot be determined).

	Person-of-subject	+	Number-of-subject	+	Tense/Aspect
<i>ek/kV</i>	1SG		∅	SG	∅ AOR
<i>na</i>	2SG				<i>pu</i> FUT
<i>et/yV/∅</i>	3SG				<i>is/s</i> PAST
<i>intV</i>	1INC.NON-SG		<i>au</i>	DU	<i>i</i> SUBJ
<i>ecrV</i>	1EXC.NON-SG		<i>taj</i>	TRL	
<i>ekV</i>	2NON-SG		<i>a</i>	PL	
<i>erV</i>	3NON-SG				

Under this hypothesis, various morphophonemic changes, particularly changes in vowel quality, would have occurred as these elements combined to form the Old Anejom preverbal particles as listed in Table 4. These morphophonemically altered forms then became compounds which, at the time Inglis was writing, were no longer transparently segmentable into a sequence of discrete meaningful units.

The Modern Anejom preverbal particles in which we are interested are listed in Table 5, and these also require some discussion before the changes over the past century are catalogued. First, a number of these particles are currently showing a tendency to become cliticised, particularly where there are no intervening particles. Most Anejom verbs are vowel initial, due to the historical accretion of an earlier morpheme, probably *a*, which was most likely a verb formative or verb marker of some kind (and which is found also in the languages of Tanna and Erromango; see Lynch to appear). In particular, consonant-final monosyllabic particles, when combining with vowel-initial verbs, show a tendency towards reduction to a single consonant proclitic; for example, *is atgii* (PAST kill.TRANS)

<sup>5</sup>Capell (n.d.) extrapolates this form on the basis of the other past trial forms.

<sup>6</sup>Inglis (1882) gives *intis* (which is formally identical to the dual form), while Capell (n.d.) gives *imjis*.

<sup>7</sup>Possibly a typographical error: *eku* is more likely.

'he killed it' becomes *s-atgii*. This reduction is optional, and appears to occur mainly in fast speech at the present time. However, the change from free particle to proclitic is one which definitely seems to be in process, and one which also probably occurred in the related languages of Tanna and Erromango some time in the past.

TABLE 5: MODERN ANEJOM PREVERBAL PARTICLES

	Singular		Non-singular	
Aorist				
1INC			<i>ekra &gt; era;</i>	<i>rai-</i>
1EXC	<i>ek;</i>	<i>k-</i>	<i>ekra &gt; era;</i>	<i>rai-</i>
2	<i>nei;</i>	<i>na-</i>	<i>ekra &gt; era;</i>	<i>rai-</i>
3	<i>et;</i>	<i>t-</i>	<i>ekra &gt; era;</i>	<i>rai-</i>
Future				
1INC			<i>ekrupu &gt; erupu;</i>	<i>rupu-</i>
1EXC	<i>ekpu</i>		<i>ekrupu &gt; erupu;</i>	<i>rupu-</i>
2	<i>napu</i>		<i>ekrupu &gt; erupu;</i>	<i>rupu-</i>
3	<i>etpu</i>		<i>ekrupu &gt; erupu;</i>	<i>rupu-</i>
Past				
1INC			<i>kis &gt; is;</i>	<i>s-</i>
1EXC	<i>kis &gt; is;</i>	<i>s-</i>	<i>is;</i>	<i>s-</i>
2	<i>is;</i>	<i>s-</i>	<i>ekris &gt; is;</i>	<i>s-</i>
3	<i>is;</i>	<i>s-</i>	<i>ekris &gt; is;</i>	<i>s-</i>
Subjunctive				
1INC			<i>ri</i>	
1EXC	<i>ki</i>		<i>ri</i>	
2	<i>ni</i>		<i>ri, ra</i>	
3	<i>iñiyi, yi</i>		<i>ri, ra</i>	

A second feature of these particles in Modern Anejom which should be mentioned here is their variability. In a number of cases in Table 5, two forms are given on either side of an arrow (e.g. *ekra > era*). In each such case, the form on the left of the arrow was elicited from an older informant in the mid 1970s while the form on the right was elicited from a younger informant six years later. Further research will be needed to establish whether the arrow represents the direction of a change or whether it merely represents generally acceptable variation.

A comparison of Tables 4 and 5 shows the changes which have taken place in the subject/tense/mood-marking system in the past century. The most obvious of these, although not perhaps the most interesting, is the loss of the distinction between dual, trial and plural numbers, with the modern language showing a distinction only between singular and non-singular. It should, however, be noted that the dual/trial/plural distinction is still maintained in the pronouns, and my most recent (and youngest) informant showed no tendency to level these distinctions in the free/focal, objective or possessive pronominal forms. It seems fairly clear that the levelling of the non-singular number distinctions in the preverbal particles has taken place simply through the loss of the dual and trial forms, with the modern non-singular forms being, in most cases, derivable from the older plural forms.

A second major change has been the tendency to level the person distinctions in the non-singular. This has gone farthest in the aorist and future, where there is no distinction between persons at all. In the subjunctive, the levelling is almost complete, although there is variability between *ri* and *ra* in the second and third persons. The past tense presents a more complex situation, for although the levelling of person distinctions in the non-singular is the least complete of all the changes in tense/aspect marking, it is almost complete in the singular; yet levelling of person distinctions in the singular of the other tense/aspect markers has not taken place at all. In fact, if the current tendency continues, there will soon be no person or number distinctions in the past tense at all. It is thus possible that the developing past tense/aspect marker *is* will come to function like a straightforward aspect marker (such as *n̄an* and *wut* in the examples given above), and may thus come to be preceded by the aorist particles; for example, Anejom may develop combinations such as *ek is*, *et is*, with the wheel turning full circle.

Not a great deal of change has taken place in the forms of the *singular* preverbal particles. In the aorist, the only change involves the second person singular, where earlier *na* is continued as the prefix *na-*, but the unprefixated form has changed to *nei*. In the past, first person *kis* is giving way to *is* as a generalised form, while second person *as* has already done so. In the subjunctive, second person *an* has become *ni*, presumably due to some kind of paradigmatic levelling. Third person Old Anejom subjunctive *inyi* may in fact be phonologically identical to Modern Anejom *iñiyi*, but it has also developed a short form *yi*.

The Modern Anejom *non-singular* forms – or at least those to the right of the arrow where there is variation – tend to continue or derive from the old third person plural forms: aorist *era* from *era*, future *erupu* from *erupu*, subjunctive *ri* from *ari*. In the past, however, the non-singular form *is* derives from the third person *singular* form *is*; this, as we have already seen, seems to be a different process. I am at a loss to explain the forms to the left of the arrow in Table 5: forms such as *ekra*, *ekrupu* and *ekris* contain a /k/ which was not present in any of the earlier forms, the nearest possible source being the velar fricative /c/ of the first person exclusive forms *ecra*, *ecrupu* and *ecris*. Further investigation is needed here to determine the use of these forms.

In summary, then, the original system of preverbal particles marking person and number of the subject as well as tense and mood seems to have broken down considerably. The distinction between the three non-singular numbers has been lost, and the distinction between the four persons has almost disappeared in the non-singular. The particles involved in marking past tense have undergone further changes, in that the person distinctions are being lost in the singular as well, and it appears that neither person nor number will be marked in the past tense in the near future. The system seems destined to end up with person and number distinctions in the aorist and subjunctive only, and with these distinctions embracing only first singular, second singular, third singular and non-singular. The morphemes *pu* 'future' and *is* 'past' are being treated, or are tending towards being treated, as aspect particles only. Finally, the system of *particles* is tending to give way to a system of *proclitics*.

## 6. DISCUSSION

The lexical changes discussed in section 2 are similar in many ways to the changes which have taken place in Mele-Fila, as discussed by Clark (1982). He could find no satisfactory explanation for the 'unnecessary' borrowings in Mele-Fila, and nor can I for Anejom. What we can say is that not only the simple fact of missionisation and European contact, but also the quite severe depopulation



which took place in the first century of that contact,<sup>8</sup> must have caused Anejom society to become very unstable indeed. In the context of that instability, it is perhaps remarkable that so little 'unnecessary' borrowing from a prestige language has taken place. The fact that Aneityum was (almost certainly) a monolingual island up until the time of contact may possibly have something to do with that.

Similarly, I also find it difficult both to explain the systemic changes in the numeral system (5>4, 6>5) and, more significantly, to account for the coincidence – if coincidence it is – of exactly the same systemic changes taking place about the same time in distant Ura. This certainly is an area which requires further investigation in other languages which have undergone drastic change in the last century or so.

Turning now to the changes in the verbal system which I discussed in the previous section, what is remarkable is not so much the loss of the distinction among the non-singular numbers but the fact that this distinction has only been lost in the system of preverbal particles and not in other forms of a pronominal nature. The Anejom pronoun system itself appears to have undergone no significant change at all since the publication of Inglis's grammar, and Anejom-speakers still preserve the four persons and four numbers in the free/focal pronouns, in the objective pronominal suffixes to verbs and in the possessive pronominal suffixes to nouns and possessive morphemes (cf. Lynch 1982a). The loss of these distinctions in the verb phrase thus requires comment.

While we cannot give an *explanation* of *why* the change took place as it did, we *can* explain why the change has not been radically disruptive, that is, why it has been relatively easily accommodated. As I have already mentioned, the normal phrase order in Anejom clauses is V(O)S. For example:

*Is atgii pikad a natañañ.*  
 PAST kill.TR pig SUBJ-MKR man  
 The man killed the pig.

When the subject noun phrase is a pronoun, it is normal for that pronoun to occur in the subject slot:<sup>9</sup>

*Ek pu atgii pikad añak.*  
 1SG.AOR FUT kill.TR pig I  
 I will kill the pig.

There is thus a certain amount of redundancy in Anejom clauses: information regarding the person and number of the subject may be given twice – once in the preverbal particle, and again in the subject noun phrase.

It is in this context that we can understand how the reduction in the marking of person and number distinctions in the preverbal particles became acceptable, even if we cannot tell how it was originally motivated. The somewhat unwieldy system of Old Anejom preverbal particles contained a number of

<sup>8</sup>See, for example, McArthur and Yaxley (1968), McArthur (1974) and Spriggs (1981) for a discussion of the nature, degree and causes of depopulation on Aneityum in particular and the southern islands in general.

<sup>9</sup>All free/focal pronouns in Anejom begin with /a/, which is probably originally the subject-marker *a*. However, although the subject-marker *a* precedes a noun phrase subject, it does not precede a pronoun subject.

redundancies, since subject pronouns and subject noun phrases also occurred postverbally. The loss of the person and number-marking functions of the preverbal particles did not place any greater decoding burden on the hearer, since the necessary information occurred later in the clause. Thus although we cannot explain *why* the change took place, we can see *how* it was easily accommodated into the existing structure.

## BIBLIOGRAPHY

- CAPELL, A., n.d., *A re-study of the language of Aneityum, New Hebrides*. Typescript.
- CLARK, Ross, 1982, 'Necessary' and 'unnecessary' borrowing. In Amran Halim, Lois Carrington and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, vol. 3:137-143. *PL*, C-76.
- CODRINGTON, R.H., 1885, *The Melanesian languages*. Oxford: Clarendon Press.
- GABELENTZ, H.C. von der, 1861-73, *Die melanesischen Sprachen*. Leipzig: S. Hirzel.
- HEWITT, Helen-Jo Jakusz, 1966, Aneityum of the Southern New Hebrides: Anejom segmental phonology and word list – a preliminary report. *Te Reo* 9:1-43.
- INGLIS, John, 1882, *A dictionary of the Aneityumese language*. London: Williams & Norgate.
- KERN, Hendrik, 1906, *Taalvergelijkende verhandelingen over het Aneityumsch, met een aanhangsel over het klankstelsel van het Eromanga*. Verhandelingen van de Koninklijke Nederlandsche Akademie van Wetenschappen, afdeling Letterkunde 8. Amsterdam: Muller.
- LYNCH, John, 1976, Notes on Aneityumese grammar. University of Papua New Guinea: mimeographed.
- 1982a, Anejom grammar sketch. In John Lynch, ed. *Papers in linguistics of Melanesia, No.4*:93-154. *PL*, A-64.
- 1982b, [Anejom lexical file]. Computer printout.
- 1983, Ura grammar sketch and vocabulary. In John Lynch, ed. *Studies in the languages of Erromango*, 148-183. *PL*, C-79.
- to appear, Verb-initial vowel accretion in Southern Vanuatu languages. In T.E. Dutton, M.D. Ross, and D.T. Tryon, eds *The language game: papers in memory of Donald C. Laycock*. *PL*, C-110.
- MCARTHUR, Norma, 1974, Population and prehistory: the late phase on Aneityum. PhD dissertation, Australian National University.
- MCARTHUR, Norma and J.F. YAXLEY, 1968, *Condominium of the New Hebrides: a report of the first census of the population, 1967*. Sydney: New South Wales Government Printer.
- RAY, Sidney H., 1926, *A comparative study of the Melanesian island languages*. Cambridge: University Press.
- SPRIGGS, M.J.T., 1981, Vegetable kingdoms: taro, irrigation, and Pacific prehistory. PhD dissertation, Australian National University.
- TRYON, D.T., 1976, *New Hebrides languages: an internal classification*. *PL*, C-50.

THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO  
THE UNIVERSITY OF CHICAGO

## PRONOUNS, POLITENESS AND HIERARCHY IN MALAY

RICHARD MCGINN

### 1. INTRODUCTION<sup>1</sup>

All Indonesian languages seem to display some mechanism for the linguistic expression of what Brown and Gilman (1960) call "the power semantic" (see Appendix). In the simplest case, a second person singular pronoun is exchanged asymmetrically for a special honorific pronoun (which may double as the plural form). The meaning of the two pronouns, over and above their reference, has roughly the force of Latin *tū* and *vōs*. As has become standard in the literature, I shall refer to these meanings as T and V. A well-attested case is Toba Batak, spoken in North Sumatra. In this language the asymmetrical exchange of *ho* 'intimate' and *hamu* 'honorific' expresses the power semantic independently of the rest of the linguistic and social context. Available choices generate the following relationships.

(1) Toba Batak power/solidarity semantic

*ho* + *ho* = T+T = social equality/intimacy  
*hamu* + *hamu* = V+V = social equality/mutual politeness  
*ho* + *hamu* = V+T = inequality/power

To be sure, numerous overtones accompany the use of pronouns in T/V languages like Toba Batak and the European examples discussed by Brown and Gilman: gestures, tone of voice, willingness to obey – and a host of supplementary linguistic forms like vocatives, titles, kin terms, names. These may add to or subtract from the status accorded an addressee by pronoun-choice. But supplementary messages can be avoided; the pronouns are often compulsory.

Another mechanism for conferring status on an addressee is the well-known case of the Javanese speech-levels (Geertz 1960). According to Errington (1986) the speech-levels can profitably be analysed and understood in terms of Brown and Gilman's power/solidarity semantics. Errington argues convincingly that the following are implied by choice of speech level in Javanese.

(2) *ngoko* + *ngoko* = T+T  
*krama* + *krama* = V+V  
*ngoko* + *krama* = T+V

---

<sup>1</sup>I wish to thank Zainab Bakir, Director of the Research Institute, Sriwijaya University, Palembang, for much valuable assistance. Also to be thanked are Ellen Rafferty and Tish Bautista for extremely helpful criticisms of an earlier draft. All errors of fact and interpretation remain mine alone.

This paper will describe a third type of power/solidarity semantic found in Indonesia, that of Standard Malay.<sup>2</sup> A brief overview of available second person address-forms is provided by four example sentences below. These data are taken from a comic book and the December 1989 issue of *Bobo*, a magazine for children published in Jakarta. They illustrate three high-frequency dyads, namely, adult-adult, mother-child and father-child. (In the examples all second person references will be presented in capital letters.)

In the first example, an adult comic-book character (who happens to be a duck) speaks to a Police Inspector (a dog) who has burst in saying "May we come in?" The answer:

- (3) *ANDA sudah di dalam.*  
 YOU already in(side)  
 YOU are already in. *Menjadi Sinterklas AWD.424:32*

In the next example, a mother speaks to her seven-year-old daughter:

- (4) *KAU tidak ada pekerjaan rumah?*  
 YOU not have work-home  
 Don't YOU have homework? *Bobo 37.XVII:7*

In the third example, a father addresses his daughter:

- (5) *IRA pasti malu memakai jas hujan ini.*  
 IRA is ashamed wear coat-rain this  
 YOU are embarrassed to wear this raincoat. *Bobo 36.XVII:33*

In the fourth example, the daughter responds to (5):

- (6) *Berkat jas hujan AYAH, Ira<sup>3</sup> tidak terlambat.*  
 Thanks coat-rain FATHER Ira not late  
 Thanks to YOUR raincoat, I was not late. *Bobo 36.XVII:33*

The problem is to account for the fact that the analogue of English 'you (singular)' is expressed in four ways in the four examples: in (3) the pronoun *ANDA* occurs; in (4) the pronoun *KAU* is used; in (5) the name *IRA* appears; and in (6) the kin term *AYAH* occurs. As the paper will demonstrate, the structure underlying these choices is quite complex and determined by a suitable generalisation of Brown and Gilman's rules. For example, the exchange of a name in (5) for a kin term in (6) will be accounted for in terms of a presumed asymmetrical exchange of T and V between father and daughter, both of whom speak in a speech-'mode' distinguished by the total avoidance of first and second person pronouns. In contrast, the choices of *ANDA* and *KAU* in (3) and (4) represent a second class of asymmetrical exchange wherein T and V are expressed by selection of an appropriate pronoun (*KAU* or *ANDA*).

The obligation to choose between T and V is the distinguishing feature of T/V languages. The data of examples (3)-(6) seem to support the idea that Malay is a special kind of T/V language. This claim constitutes my working hypothesis. Nevertheless, Malay obviously differs from 'classic' T/V languages because it has the ability to confer relative status on an addressee by means other than

<sup>2</sup>'Standard Malay' is an idealisation I will use throughout the paper to refer to common structural features shared by the national languages of Indonesia and Malaysia. However, the actual data I will cite is restricted to written sources originating in Indonesia.

<sup>3</sup>Ira uses a pronoun-substitute as first-person reference instead of the pronoun *saya*. This feature is explained later in the paper.



pronouns. In fact, the device of using kin terms and names as second person address forms means it has (virtually) an open class of distinctions at its disposal.

An obvious implication of the hypothesis may be derived from Brown and Gilman's rules. T/V languages seem indicative of ideology; in particular, they are linked to 'stratified' societies. Does the implication apply in the case of Malay? Much recent literature suggests that Malay would be considered an exception to this generalisation. For example, Malay is often characterised as 'flat' and relatively egalitarian compared to languages like Toba Batak and Javanese. The supposed 'relative stylistic simplicity' of Standard Malay is praised as a special political and social advantage.<sup>4</sup> Colourlessness is what is sought; a national language ought to be egalitarian and offer relief from the pressures of linguistic etiquette imposed by the 'native' languages. This theme has been sounded recently by Lowenberg (1990:112), who quotes B. Anderson (1966) concerning the political role played by Malay in the early part of this century:

It [Malay] was a language simple and flexible enough to be rapidly developed into a modern political language...This was possible because Malay had *ipso facto* an almost statusless character, like Esperanto, and was tied to no particular regional social structure. It had thus a free, almost 'democratic' character from the outset...(Anderson 1966:104)

Claims like this are common enough in the literature, but they are restricted to broad summaries and are not backed up by analysis. In contrast, analytical studies of Malay which might be offered as evidence for or against such claims do not pay attention to broader implications. As far as I know, there are no analytical studies of Malay that have claimed that the language is particularly simple, nor indeed statusless and free of ties to social structure. So there is a discrepancy between what one is led to believe about the language by reading a typical summary statement about the status of Malay on the one hand, and a typical analysis of Malay structure on the other.

This paper will attempt to bridge the gap by arguing that broad, often-quoted claims like the above reflect an 'outsider's view' of Malay that rests on a very insecure foundation. It is a view that seems to have developed in the minds of foreign scholars and other second-language users of Malay (including Indonesians) who for one reason or another have felt obliged to give reasons why Malay has proven to be such a felicitous choice to serve as the basis of the national languages of Indonesia and Malaysia. The purpose of this paper is to do a small bit of analysis and then to point out why the analysis (and by implication any serious analysis of the language) fails to support the standard view.

An interesting case in point concerns the recent introduction of the second person pronoun *anda* in both Indonesia and Malaysia. According to Kridalaksana (1974:20, n.5), *anda* 'you (singular)' was coined by governmental language-planners in 1957 to eliminate "the chaotic, undemocratic and inefficient" Indonesian pronominal system. Kridalaksana remarks that whereas *anda* "has become widespread and has certainly enriched the vocabulary, it is hard to say that it has made the Indonesian system more 'democratic' than it was".<sup>5</sup>

This paper will describe the 'undemocratic' pronominal system of Standard Malay as anything but chaotic and inefficient. The paper will argue that (a) Malay displays a richer power semantic than

<sup>4</sup>Errington (1986:335) notes that Malay's perceived simplicity made it 'a congenial choice of the nationalist movement' in Indonesia.

<sup>5</sup>Kridalaksana (1974) cites the 1957-58 volumes of the journal *Bahasa dan Budaya* for discussion of this subject. See also Prentice (1987:930), who states that *anda* has only added to the complexity of the Indonesian pronoun system.



classic T/V languages like Latin and Toba Batak, and that (b) the structure of the power semantic in Malay probably is similar – and may even be identical – to that of Javanese.<sup>6</sup>

To anticipate briefly, I will look at some evidence suggesting that second person address forms in contemporary Standard Malay are heavily power-laden in the sense of Brown and Gilman's rules. I then try to draw some sociolinguistic implications. In particular, the facts seem to refute the claim that Malay is 'democratic' and 'free of ties to social structure'. Furthermore, I will argue that the presence or absence of supposedly undesirable linguistic features may not directly bear on the issue of a language's political role. What counts in the political domain is not linguistic structure but history – the weight of years and the numbers of second-language users.

I wish to avoid the impression that the argument will proceed from an a priori premise. I believe specific evidence can be adduced that leads to the conclusion I aim to draw, namely, that obligatory linguistic features that carry implications of social hierarchy are compatible with the language's wider political and social role.

Among the things I shall take for granted in this paper are the following three: (a) there are native speakers of Standard Malay; (b) Malay speakers are comfortable using formal and informal styles; and (c) the public domain abounds with examples of Standard Malay in written form – newspapers, television, magazines, books (including comic books). In the public domain, Standard Malay is called Bahasa Indonesia (in Indonesia) and Bahasa Malaysia (in Malaysia).

Before concluding this introductory section, it is perhaps useful to make explicit an assumption about the notion 'honorific pronoun' or 'pronoun used to give V' that will be used in the paper. I assume that the following two formal features define an honorific pronoun. First, to be considered an honorific pronoun the form must refer unambiguously to second person *singular*, and be distinct from at least another second person singular pronoun that is used non-honorifically. These two criteria exclude 'you (singular)' in English, and they also exclude Malay *kamu*, which does not contrast with *kau* in this way.<sup>7</sup>

The criteria also exclude forms such as *IRA* and *AYAH* used as 'pronoun-substitutes' in examples (5) and (6). This term is adopted from Dardjowidjojo (1978), who used it in relation to Javanese. Pronoun-substitutes are used in Malay (as well as Javanese) as second person address forms. I will argue that *IRA* and *AYAH* in examples (5)-(6) are pronoun-substitutes that effectuate asymmetrical exchange of T and V between different-status individuals. By the above criteria, pronoun-substitutes are not simply a class of pronouns since there is inherent lexical content. The evidence for this claim is that pronoun-substitutes are systematically ambiguous whenever they occur: the actual reference (either third person or second person) must be inferred from context by the listener.<sup>8</sup>

The two formal features combine with a single pragmatic feature as well. That is, to be considered honorific, a pronoun must actually be used honorifically. The pragmatic feature rules out many dialect forms as honorifics. For example, Ujan Mas Malay (South Sumatra), and Bengkulu Malay

<sup>6</sup>This implies that the Javanese power semantic is distinguished by the fact that 'power words' beyond address forms are elaborated in the lexicon (Errington 1986).

<sup>7</sup>Historically it apparently did (Prentice 1987). Interestingly, there exist Malay dialects in South Sumatra that still use *kamu* honorifically. See Appendix.

<sup>8</sup>The criteria do not support characterisations like the following: "Indonesian is one of the few languages in the world in which pronouns are an open class with an infinite membership" (Prentice 1987:931). Rather, it is the pronoun-substitute class that is 'open' in the sense that it includes all names and kin terms.

(Bengkulu) distinguish second person masculine and feminine pronouns in the singular, and display a second person plural pronoun used for both sexes, but none of these forms are used honorifically. Rather, to confer status upon an addressee, these dialects utilise pronoun-substitutes exactly like Standard Malay. (See Appendix.)

The data for the first part of the paper have been limited to texts written specifically for Indonesian children. Later in the paper I will attempt to relate these data to some recent studies of spoken Bahasa Indonesia. Children's magazines and comic books offer an accessible, and reasonably rich, sampling of data on second person pronoun usage and avoidance. The data were taken from two series, both published in Jakarta: (a) the November and December, 1989, issues of *Bobo*, a children's magazine, and (b) five Walt Disney comic books translated into Indonesian from English, purchased in 1989. The sources are rather impoverished culturally speaking (a fact that will be emphasised when gaps in the data are discussed). For this very reason, I believe the data represent with special clarity the structure of the power semantic.<sup>9</sup> Even in the most culturally bleached of children's stories, systematic and obligatory linguistic expressions of the power semantic are found on every page.

## 2. THE STRUCTURE OF THE POWER SEMANTIC IN WRITTEN STANDARD MALAY

In the sources I examined, the asymmetrical exchange of T and V takes place in two distinct 'modes' that I shall call Distal Mode and Familial Mode. In Distal Mode, T and V are exchanged by pronouns, as in examples (3) and (4); in Familial Mode, T and V are exchanged by names and kin terms used as pronoun-substitutes,<sup>10</sup> as in examples (5) and (6). The term Distal is meant to imply 'relative social distance', which is the interpretation assigned to the V pronoun *ANDA* in this system; the T pronouns *KAU* and *KAMU* are accordingly interpreted as expressing 'social proximity' or 'intimacy'. In the Familial Mode one avoids second person pronouns and uses pronoun-substitutes in their place. To express V in the Familial Mode, an appropriate kin term is given as address-form; to express T in the Familial Mode, the addressee's name or nickname is given.

The pronoun-substitutes form a special category of nominals described by Dardjowidjojo (1978). Pronoun-substitutes are used in argument positions (subject, object, genitive) and refer to an addressee. In this paper I shall be concerned **only** with argument positions. (For an interesting study of names and titles used vocatively in Bahasa Indonesia, see Jenson 1988.)

<sup>9</sup>The content (as distinct from the language) of some of the data is devoid of references to Indonesian culture. The Disney comics, in particular, feature talking animals in settings that are either ambivalent or obviously Western. Even children's magazines like *Bobo*, although often depicting human kids in what look to me like typical Indonesian cultural situations, sometimes feature stories with decidedly 'international' themes. For example, the December 1989 issue of *Bobo* features one translated story from Holland and one comic section featuring the 'Snow Queen'.

<sup>10</sup>In the seven sources I examined, only the following pronoun-substitutes actually occurred. (See Kridalaksana 1974 for a more complete listing of pronoun-substitutes in common use in spoken Bahasa Indonesia.)

### 1st and 2nd Person Pronoun-Substitutes

Familial T (fT):     Ira  
                           Fia  
                           ... (i.e. all children's names)

Familial V (fV):  
     *Ayah*     'father'     (used to refer to one's biological father)  
     *Bapak*    'father'     (general term of respect for males)  
     *Ibu*       'mother'    (used for biological mother or as general term of respect for females)  
     *Paman*    'uncle'

The simple fact that T and V can be given in either of two modes generates a startling number of possibilities for expressing power relationships – namely ten. The dyad-types – or possible power relationships – fall into three sets as displayed in Table 1.

TABLE 1: MALAY POWER/SOLIDARITY SEMANTIC

(A) Distal Speech Mode				Example
I	T pronoun	+ T pronoun	= dT+dT	<i>KAU + KAU</i>
II	V pronoun	+ V pronoun	= dV+dV	<i>ANDA + ANDA</i>
III	V pronoun	+ T pronoun	= dV+dT	<i>ANDA + KAU</i>
(B) Familial Speech Mode				
IV*	name	+ name	= fT+fT	<i>*IRA + ALI</i>
V	kin term	+ kin term	= fV+fV	<i>IBU + BAPAK</i>
VI	kin term	+ name	= fV+fT	<i>IBU + IRA</i>
(C) Mixed-Mode Dialogue				
VII*	T pronoun	+ V name	= dT + fT	<i>*KAU + NAME</i>
VIII*	V pronoun	+ V name	= dV + fT	<i>*ANDA + NAME</i>
IX	T pronoun	+ V kin term	= dT + fV	<i>KAU + IBU</i>
X*	V pronoun	+ V kin term	= dV + fV	<i>*ANDA + IBU</i>

(Dyad-types for which no data were found in the sources are marked with an asterisk (\*) and discussed in the next sub-section of the paper.)

The ten power relationships defined in Table 1 constitute the minimum structure needed to account for the four second person address forms in examples (3)-(6) above. The remainder of this section will be devoted to justifying this claim.

The pivotal opposition, the one that expresses the maximum asymmetry, is 'Familial V' (fV) and 'Distal T' (dT). This power relationship is represented as IX in Table 1. It is doubly asymmetrical because the superior partner expresses closeness with the T pronoun *KAU* 'YOU' in exchange for Familial V (e.g. kin term *IBU* 'MOTHER'). This relationship has been observed to exist in all mother-child dyads I found in the data. I shall return to this point later in the paper.

First let us examine the dyads that brought out **pronoun** choices from both participants (see Table 1 (A) – 'Distal Mode').

#### (A) T and V in Distal Mode: dT and dV

A total of six second person pronouns occur with high frequency in the data. One is the newly-coined *ANDA*, which occurs systematically in the comics in both symmetric and asymmetric dyad-types. *ANDA* is exchanged symmetrically by adult animal characters who are not well acquainted. Recall example (3) above. An adult character had discovered a police inspector in his house; the police had just burst in saying *Boleh kami masuk?* ('May we come in?') This question is answered by (3) (*ANDA sudah di dalam* 'YOU are already in').

*ANDA* (dV) also occurs in asymmetrical exchanges between adults in exchange for *KAU* and *KAMU*. An extended example occurs in the Disney comic AWD.422:8-16 titled *Ayam Bertelur Emas*. The story features Uncle Scrooge in the role of a foreigner speaking with another older man in

the role of native, or host, in a land Scrooge is visiting. Scrooge gives *ANDA* (dV) and receives *(ENG)KA<-M>U* (dT) throughout this story.

In the comics *ANDA* (dV) contrasts with a number of Distal Mode T pronouns (dT), which imply closeness or intimacy. The dT pronouns are *KAU*, *ENGKAU*, *KAMU*, the post-clitic *-MU* (all singular) and *KALIAN* in the plural. This set of five pronouns is exchanged symmetrically in child-child dyads and also in adult-adult dyads; and is given asymmetrically by adults to children in exchange for fV. The use of *-MU* is limited to post-nominal and post-verbal positions. I shall not attempt the formidable task of describing and differentiating these T pronouns beyond singular and plural. For my purpose, it is sufficient to observe the general fact that as a set they are *invariably* given to children in the comics. For this reason I have assumed that these pronouns represent the 'lowest' power word in the hierarchy. For convenience, I shall follow E. Anderson (1983) in treating all five pronouns as equivalent. I shall henceforth represent them all by the admittedly awkward formula *(ENG)KA<- M>U*.

In the first example *(ENG)KA<-M>U* is used by a child to address his pet.

(7) Child:

*Oh, KAMU tetap mau ikut main?*  
 Oh, YOU-dT still want to play?  
 Oh, YOU-dT still want to play? *Bobo 36:XVII.17*

The following symmetrical exchange of dT occurs between a rabbit and an elephant.

(8) Rabbit:

*Aku akan segera memanggil kawan-kawan-ku untuk menolong-MU.*  
 I will soon call friend-s my to help YOU  
 I will call my friends to help YOU-dT.

(9) Elephant:

*KAU telah menyelamatkan diri-ku.*  
 YOU have save self-my  
 YOU-dT have saved me. *Bobo 36:XVII.5*

Adult animals in the comics 'think aloud' to themselves and converse with other animal characters in the same Distal Mode. Consider the following dialogue from a Walt Disney comic.

(10) Donald Duck (telling a lie to Daisy Duck):

*Sayang, saya<sup>11</sup> harus membuatnya kecewa! Karena sore*  
 too bad I must make-him disappointed because afternoon  
*ini harus mengantarkan KAMU ke tempat puisi!*  
 this must accompany YOU to place poetry  
 Too bad I'll have to disappoint him. Because (I) have to take YOU-dT to the poetry reading this afternoon.

(11) Daisy Duck's response to Donald:

*Yah, tidak apa-apa, Donal! Lebih penting kalau KAMU*  
 Oh it doesn't matter Donald more important that YOU-dT

<sup>11</sup>*Saya* occurred only once in the five comics I examined, *aku* being much the preferred form.

*menengok mantan guru-MU itu.*  
 visit former teacher-YOUR-dT the

Oh, it doesn't matter, Donald. It's more important that YOU visit YOUR former teacher.

*Polisi Gunung AWD 406.31*

To summarise second person pronoun usage, in the comics (ENG)KA<-M>U is always exchanged symmetrically between low-status individuals such as same-age children. (ENG)KA<-M>U is also exchanged symmetrically by same-age adults who seem to be friends. In contrast, ANDA is exchanged symmetrically between adults who are not acquainted (e.g. policemen, store clerks and the like); and finally ANDA may be exchanged asymmetrically for (ENG)KA<-M>U between adults of clearly different generations. In such cases, ANDA is always given by the younger adult in exchange for (ENG)KA<-M>U from the elder.

#### (B) T and V in the Familial Mode: fT and fV

In this sub-section, the pattern displayed in Table 1 (B) – the giving of T and V in the Familial Mode – will be illustrated. Familial dialogue occurs when second person pronouns are avoided by both participants, who employ pronoun-substitutes in their place. Several examples of two-way Familial Mode dialogues were found in *Bobo*; interestingly, none occurred in the five translated Disney comics I examined.<sup>12</sup>

A secondary diagnostic of the Familial Mode is replacement of *aku*<sup>13</sup> either by *saya* 'I' or the speaker's name.<sup>14</sup> Another diagnostic – and the one I am most interested in here – is the avoidance of second person pronouns. The third and final diagnostic is the selection of pronoun-substitutes in grammatical environments where second person pronouns would be expected to appear in the Distal Mode.

<sup>12</sup>I can think of no reason for this and would expect to find some examples of two-way, Familial Mode dialogues if more comics were searched.

<sup>13</sup>It seems reasonable to suggest that aspects of the Distal Mode could be characterised as 'private-language', i.e. the language Malay speakers use when thinking to themselves. This characterisation would account for the fact that Malay speakers generally report using *aku* (never *saya*) when thinking privately to themselves.

Not unexpectedly, people tend to 'think' in their native dialect. In the South, this means quite a different variety of Malay than the one that appears in the comics. But many Sumatrans from the northern provinces reported to me that they think in Standard Malay, and in a style that closely resembles the comic book style.

For what they are worth, the following anecdotes reveal some of the problems involved in any attempt to elicit 'private language' data from informants.

While travelling through Sumatra and Malaysia in November and December, 1989, I asked people from different dialect areas how they address themselves when they commit an obvious blunder. To elicit this information, I offered that in English, when angry at myself I usually address myself as 'you' or sometimes 'I' and say something like: "You idiot!" or "I must be nuts!"

Based on scattered informants' responses to this informal inquiry, I discovered that Malays do not address themselves with a second person pronoun equivalent to my 'You \_\_\_\_!'; instead, they all reported using only a first person reference. Furthermore, all reported that the only pronoun possible for them in this kind of 'private' language is *aku* 'I' (never the Standard Malay pronoun *saya*). A typical response was: *Alangkah bodohnya aku ini!* ('How stupid I am!'). Significantly, there was agreement on the choice of *aku* over *saya* in this context. Thus, it appears that *aku* may be one diagnostic for private-language pronoun, whereas *saya* is indicative of 'social discourse'. See fn. 8..

The only clear case of a second person pronoun being used in 'private language' involved prayers addressed to the Deity. Nearly all of my informants reported using *Engkau* for this purpose.

<sup>14</sup>See fn. 3.



Theoretically, two-way Familial Mode dialogues may be symmetrical (T+T or V+V) or asymmetrical (T+V). However, the sources I examined produced no examples of T+T in the Familial Mode (NAME + NAME), but there were several examples of V+V and T+V.

An example of a symmetrical Familial Mode dialogue (V+V) is the following. In the story, a middle-aged woman is being drawn into a get-rich-quick scheme. When the con man arrives at the door the woman greets him politely:

(12) Adult woman to man at door (fV):

*Ee ... Pak Jonto. Mari, Pak, silakan duduk di dalam.*  
 Ee ... Mr Jonto come sir please sit in(side)  
 Ee...Mr Jonto. Come in, Sir, please sit down inside.

(13) Man at the door (fV):

*Tak usah repot-repot, Bu. Saya nanti akan ke sini sekitar jam 6.*  
 no need trouble Ma'am I later will to here around o'clock 6  
 No need to bother, Ma'am. I will be back here later at about 6 o'clock.

(14) Woman's response (fV):

*Baiklah, saya tunggu kedatangan BAPAK.*  
 fine I wait arrival YOUR  
 Fine, I will wait for YOUR-fV arrival. *Bobo 37.XVII:33*

In this dialogue, the woman and man exchange V in the Familial Mode (fV + fV). As an indication of this, both use the first person pronoun *saya*; the vocative positions are filled by honorific words *Bu* and *Pak*.<sup>15</sup> Most importantly for the purposes of this paper, in example (14) the woman uses the second person pronoun-substitute *BAPAK* 'father' in genitive position.

Asymmetrical (fT + fV) exchanges may also occur entirely within the Familial Mode. This type of dyad was illustrated in example (5) and (6) above. Another example features the same father-daughter dyad. The daughter is named Ira. Judging from the illustrations, Ira is about seven years old. When speaking to her father, Ira gives fV and receives fT. An added twist is that both participants avoid the first person pronoun (*saya*) as well as second person pronouns. In other words, all references to either addressee are made with pronoun-substitutes.

(In the following examples, first person references will be underlined and second person references appear in capital letters as before.)

(15) Daughter to father (fV):

Ira tidak mau pergi ke sekolah.  
Ira not want go to school  
 I don't want to go to school.

(16) Father to daughter (fT):

Ayah mengerti. IRA pasti malu memakai jas hujan ini.  
father understand IRA is ashamed wear coat-rain this  
 I understand. YOU-fT are embarrassed to wear this raincoat.

That evening Father asks Ira if she got to school all right and Ira responds as follows.

<sup>15</sup>See Jensen (1988) for discussion of vocatives in Bahasa Indonesia.



## (17) Daughter to father (fV):

*Berkat jas hujan AYAH, Ira tidak terlambat.*  
 Thanks coat-rain FATHER Ira not late  
 Thanks to YOUR-fV raincoat, I was not late. *Bobo 36.XVII:33*

These choices are presumably accounted for by our working hypothesis. That is, it seems that Malay does in fact provide the necessary structure for exchanging T and V even in the total absence of second person pronouns. Assuming that a NAME (=‘YOU’) in exchange for a KIN TERM (=‘YOU’) constitutes an asymmetrical exchange of power-laden terms, it makes sense to suggest that the KIN TERM ‘means’ V and the NAME ‘means’ T. In other words, the choices made by Ira and her father are accounted for by Brown and Gilman's rules for T/V languages. The evidence supports the claim that Malay is a special kind of T/V language, one that has found a way to exchange T and V asymmetrically without pronouns.<sup>16</sup>

## (C) T and V in doubly-asymmetrical (‘Mixed’) Modes

In the sources I examined, children, when speaking to adults, seemed limited to the giving of fV (=an appropriate kin term), never dV (*ANDA*). However, a contrast was observed in the manner that adults return T to children. For example, in (5) and (16) above, a father gives the child's NAME (fT) whereas in (4) a mother gives *KAU* (dT). The latter pattern (dT for fV) was repeated in all the mother-child dyads that occurred in the sources.<sup>17</sup> This seems to constitute the polar relationship in the system (see Table 2, IX).

Again, we can turn to the Disney comic books for an abundance of examples. When speaking with his three young nephews, Donald Duck, as the adult, always gives (*ENG*)*KA<M>U* (dT) and always receives the kin term *PAMAN* (fV) ‘UNCLE’ in return. The following are typical examples.

## (18) Donald to nephews (dT):

*Ayo, KALIAN masuk kamar tidur saja! Biarkan aku sendirian!*  
 now YOU-PL go room sleep only let me alone  
 Now YOU-dT go to bed! Leave me alone! *Polisi Gunung AWD 406.4*

## (19) Nephew to Donald (fV):

*Tidak datang-kah surat yang PAMAN tunggu itu?*  
 not come-question letter that UNCLE wait the  
 Didn't the letter that YOU-fV are waiting for arrive? *Polisi Gunung AWD 406.4*

## (20) Another Nephew to Donald:

*Mengapa PAMAN tidak berlibur saja dengan kami?*  
 why UNCLE not vacation only with us (exclusive)  
 Why don't YOU-fV take a vacation with us? *Polisi Gunung AWD 406.5*

<sup>16</sup>The reader can gain an appreciation for the force of the exchange of T and V by imagining Father's reaction if Ira were to address him by name instead of the honorific kin term *Ayah*.

<sup>17</sup>There were an insufficient number of father-child dyads to be able to comment on the difference between mother-child and father-child dyads. Note, however, that in terms of the working hypothesis, all parent-child dyads were asymmetrical T + V.

Another example is from the story 'Father's raincoat' cited earlier. After receiving *IBU* (fV) 'MOTHER' from Ira, mother returns (*ENG*)KA<-M>U (dT).

(21) Mother to daughter (dT):

*Nah anak yang manis,*<sup>18</sup> *gantilah pakaian-MU.*  
 Ah, child sweet change clothes-YOUR  
 Now, sweet child, change YOUR-dT clothes. *Bobo* 36.XVII:33

(D) Mode-switching

The next two examples below illustrate another twist, namely, Mode-switching by one partner in a dialogue. A switch is recognised when one partner begins in one mode (Familial) and then switches to the other (Distal). Brown and Gilman (1960:262) cite cases of speakers of European T/V languages switching from V to T in the course of a relationship or even a single conversation. Likewise, in Malay, speakers ought to be able to switch, potentially, between V and T and also between Familial Mode and Distal Mode. Interestingly, I found both types. Most interesting were the switches from Familial to Distal Mode (maintaining V as a constant). One example involves Donald Duck (an adult) and Uncle Scrooge (a generation older than Donald – Scrooge walks with a cane). In the opening frame of one story, Donald gives Uncle Scrooge the appropriate kin term *PAMAN* 'UNCLE' (fV), but in subsequent frames Donald switches to *ANDA* (dV). For his part, the older man (Scrooge) gives only (*ENG*)KA<-M>U (dT).

As the story opens Scrooge is sitting on a huge pile of money:

(22) Donald to Scrooge (fV):

*Bagaimana perasaan PAMAN duduk di atas uang PAMAN itu?*  
 how feeling UNCLE sit on top money UNCLE that  
 How does it feel for YOU-fV to sit on top of YOUR-fV money?

(23) Scrooge to Donald (dT):

*Cemas! Sedih! Sengsara! KAU sih tenang-tenang saja, Donal...*  
 awful sad miserable YOU particle peace-of-mind just Donald...  
 Awful! Unhappy! Miserable! YOU-dT have nothing to worry about, Donald ...

In the next frame, Donald is clearly being ironic. He switches to the Distal Mode, maintaining V with the pronoun *ANDA* (dV).

(24) Donald to Scrooge (dV):

*ANDA memang pantas bersedih, yaman Gober!*  
 YOU really appropriate be-sad uncle Gober  
 YOU-dV have every right to be unhappy, Uncle Scrooge.  
*Gudang Uang Tembus Pandang* AWD 417.11

Apart from the opening frame, therefore, this story illustrates asymmetry in the Distal Mode since *ANDA* (dV) is exchanged for (*ENG*)KA<-M>U (dT).

<sup>18</sup>Both Mother and Father refer to Ira as *anak yang manis* in vocative position, but this expression does not occur in argument position.

## (E) Gaps in the data

Of the ten possible power relationships displayed in Table 1, only six were actually found in the data (see Table 1). The power relationships observed, together with the gaps in the data, are listed again and displayed alongside information about the speakers in Table 2. Roman numerals refer to the same dyad-types listed in Table 1.

TABLE 2: GAPS IN THE DATA

Potential dyad types				Evidence for in children's magazines
(A) Distal Mode				
I	(ENG)KA<-M>U	+ (ENG)KA<-M>U	= dT+dT	friend + friend (any age)
II	ANDA	+ ANDA	= dV+dV	store clerk + adult
III	ANDA	+ (ENG)KA<-M>U	= dV+dT	Scrooge + very old man
(B) Familial Mode				
IV*	NAME	+ NAME	= fT+fT	---
V	kin term	+ kin term	= fV+fV	woman + man at door
VI	kin term	+ NAME	= fV+fT	father + young child
(C) Mixed Mode				
VII*	T (ENG)KA<-M>U	+ NAME	= dT+fT	---
VIII	V ANDA	+ NAME	= dV+fT	---
IX	T (ENG)KA<-M>U	+ kin term	= dT+fV	mother + young child
X*	V ANDA	+ kin term	= dV+fV	---

Each of the four gaps involves a NAME, ANDA or both. In the sources I examined, the NAME was given only in the father-daughter dyad (examples (6) and (16) above). However, I suspect that the three gaps involving NAMEs may be accidental owing to the limited number of children's magazines I examined. Based on E. Anderson's (1983) statistical study of forms of address in Bahasa Indonesia, this would seem a reasonable guess. Perhaps further research would fill dyad-type X as well, that is, ANDA + KIN TERM (Anderson does not mention ANDA at all). All that can be said with assurance in this paper is that, in the limited sources I examined, the gaps are simply unexplained.<sup>19</sup>

Details aside, however, the conclusion seems unaffected by these gaps in the data. Malay's system of second person address forms is structured to provide ten distinct dyad-types. Even if some of the potential dyad-types are unrealised, it is not possible to simplify the structure and still account for the four categories of second person address forms that occur regularly in the data. The system conforms to the rules and expectations of Brown and Gilman's broad study of T/V languages. More than that, it does so not once but **twice**: in Distal Mode and in Familial Mode.

<sup>19</sup>Also unattested in my data are combinations of title+name used as pronoun-substitutes, e.g. Mas Mus, Mbak Tati, Bu Bandi, Pak Tikno and so on. The literature cites these as high-frequency combinations in spoken Malay. E. Anderson (1983) gives examples and percentage of use of these combinations in comparison with the occurrence of (ENG)KA<-M>U and kin terms. His paper will be discussed in detail in the next section.

### 3. IMPLICATIONS FOR THE STUDY OF SOCIAL HIERARCHY

In their most interesting generalisation, Brown and Gilman (1960) suggest that there is a worldwide trend towards egalitarianism in T/V languages that is marked by an increased tendency to ignore status differences. This means an increased tendency for family members to exchange T among themselves regardless of age and status; and for adult non-family non-intimates to exchange V symmetrically.

Both kinds of trend toward equality (and away from the power-laden asymmetry implied by T + V) are triggered internally by the 'tension' that is inherent in T/V languages. Brown and Gilman (p.258) point out that:

The dimension of solidarity is potentially applicable to all persons addressed. Power superiors may be solidary (parents, elder siblings) or not solidary (officials whom one seldom sees). Power inferiors, similarly, may be as solidary as the old family retainer and as remote as the waiter in a strange restaurant.

Does Malay conform to the world-wide tendency towards democratic speech? Do 'modern' Malay-speaking parents increasingly draw T from their children nowadays? I found no evidence of this in the children's sources I examined. All parent-child dyads in the data were asymmetrical (T + V).

This fact might have significance beyond the data. Recall that some of my data involved 'fantastic' animal characters. If an 'egalitarian' trend in child-parent relationships were really a fact of spoken Standard Malay, one might expect the trend to show itself in these culturally-neutral settings. But no such trend was found. Comic-book animal-parents **never** exchanged solidary T or V with children; they **invariably** drew V and gave T. Further questions to ask are: How do 'real' modern parents relate to their grown children? Grown older-siblings to grown younger-siblings? Husbands to wives? Is there any evidence that T + T is on the rise among adult-intimates in Modern Indonesia or Malaysia? I shall return to these questions directly below.

Among adult non-intimates, on the other hand, V + V was common in the data (in both Modes). This fact raises further questions concerning its relevance to spoken Malay viewed now as a vehicle of inter-ethnic communication outside the home. Do adult-adult dialogues involving non-intimates tend to be 'more democratic' nowadays? In particular, are exchanges of *ANDA* + *ANDA* or *KIN TERM* + *KIN TERM* common between boss and employee? Officer and soldier? Passenger and petycab driver? And so on.

As a beginning towards answering these kinds of questions, E. Anderson's (1983) statistical study of linguistic variation in Bahasa Indonesia as spoken in Bandung, West Java, offers some interesting data for interpretation. Anderson reported no instances of *ANDA* in his sample. Furthermore, he recorded no evidence of *(ENG)KA<-M>U* exchanged symmetrically (T + T) between same-sex **strangers**, and also none between same-sex **co-workers**. Between same-sex **friends** Anderson recorded only one case of *(ENG)KA<-M>U* (this constituted only 2% of the sample); whereas among **family members** he recorded only four instances of *(ENG)KA<-M>U* or 13% of the sample – all confined to the speech of young males. In sharp contrast, same-sex **strangers** used *KIN TERMS* 100% of the time, and same-sex **co-workers** and **friends** also used *KIN TERMS* 95% of the time. About half the cases displayed the *KIN TERM* with the *NAME* attached (e.g. *PAK SUTEDJO* in place of unmodified kin term *BAPAK*). I have treated this variation as insignificant (both signalling fV) for the purposes of comparison with the data from the children's magazines.

In sharp contrast, again, unmodified NAME (without a kin term attached = T) was rarely used outside the family: 5% between co-workers and 2% between friends; but the addressee's NAME was used relatively frequently inside the family (27%). In fact, if I interpret Anderson's data and analysis correctly,<sup>20</sup> the major interactive device within parent-child dyads involved the parent giving the child's NAME and drawing the appropriate KIN TERM. If so, Anderson's study clearly supports the hypothesis of this paper that T and V not only exist in Malay, but that T and V can be exchanged entirely in Familial Mode (using pronoun-substitutes instead of pronouns). Anderson's study thus validates the dialogue between Father and Ira recorded as examples (5)-(6) of this paper, and the interpretation that was given there.

Most of Anderson's figures can be interpreted in terms of the power-solidarity semantic and Brown and Gilman's rules. However, some interesting questions remain unanswered, especially with respect to his data on differences between men's and women's speech.<sup>21</sup> Anderson found that women tended to speak more formally than men, that is, were overall more likely to give V to an addressee. In fact, as mentioned above, the 'lowest' address form (ENG)KA<-M>U was restricted to the speech of younger men inside the family, where it accounted for 15% of the sample. This carries the implication that (ENG)KA<-M>U is never given to an adult of either sex (a claim that would require more data to fully substantiate).<sup>22</sup>

Anderson's study, while not pretending to be a large-scale one, draws upon a richer data base than my own analysis of pronoun use and avoidance in the comics. Combining our results, it seems to me that several preliminary conclusions can be drawn which suggest the direction that future research on these topics might take.

Despite a necessary qualification concerning the observed preference for 'Familial' Mode, Anderson's study supports my hypothesis that T and V are commonly exchanged asymmetrically in spoken Malay. The evidence thus clearly contradicts the notion that Malay is 'flat' if by this it is implied that Malay lacks this capability. However, it must be acknowledged that Anderson's study does support the idea that the style of Indonesian exchanged among adults **outside the family** tends toward polite exchange of solidary V. This finding, if truly generalisable, would suggest an 'egalitarian trend' of a potentially significant kind, perhaps comparable to Brown and Gilman's international trend cited earlier in the paper. For example, according to Brown and Gilman (p.257), after the break-up of the Roman Empire:

Europeans became very conscious of the extensive use of V as a mark of elegance. In the drama of seventeenth century France the nobility and bourgeoisie almost always address one another as V. This is true even of husband and wife, of lovers, and of parent and child if the child is adult. Servants and peasantry, however, regularly used T among themselves.

<sup>20</sup>See especially E. Anderson (1983:14) for the first person and second person data displayed there.

<sup>21</sup>Another question relates to the phenomenon of switching between T and V. Anderson reported no cases of this. If such were found, it would constitute a clear sign of instability and potential change.

<sup>22</sup>This implication is more plausible within the context of some chilling facts of recent history. Prior to the failed coup attempt in 1965, a hallmark of the Communist movement was its attempt to replace the power semantic (as a feudal remnant) by solidary T. I am told that this attempt was misunderstood by the ordinary people, who interpreted it to mean that Communists 'have no respect'. This attitude makes it unlikely that Bahasa Indonesia will soon follow the example of French in conformity with Brown and Gilman's worldwide trend toward egalitarianism, at least not in the form of symmetrical exchange of solidary T.



It seems entirely reasonable to suggest (as a working hypothesis to guide further research) that a similar trend towards 'elegance' or 'refined speech' might be the motivating force behind the extensive and symmetrical exchange of KIN TERMS (V + V) among educated persons in Indonesia. However, the mere fact that Familial Mode V + V predominates in adult-adult dyads outside the home offers insufficient grounds to conclude that the Malay language is free of obligatory references to social hierarchy, for several reasons. First, the exchange of solidary V in the form of KIN TERM nonetheless implies its own inherent asymmetry because kin terms are exchanged not only as category-types but at the same time as term-tokens. Clearly the lexical content of *NENEK* 'grandmother' is not the same as *IBU* 'mother' although both may be used categorically as second person pronoun-substitutes. In the same way, the exchange of *BAPAK* for *IBU* is potentially asymmetrical. It should not be taken for granted that the two are perfectly status-equivalent; rather, this should constitute a research question. For example, why (in most areas) do adult males draw a distinction between *AYAH* (biological father) and *BAPAK* (general respect term for males) – both translated into English as 'father' – whereas females draw only the single term *IBU* 'mother'?

Second, all the evidence suggests that explicit, obligatory reference to an addressee's relative status is a feature of stability in the historical relationship between parents and small children. This established, however, there remain open many questions (such as those posed above) that should be investigated in light of Malay's power/solidarity semantic.

Finally, my own brief analysis of address forms in the comics would seem to be most relevant for the style of Malay that is available at the other end of the social ladder – at the level of 'servants and peasantry' perhaps, of children, of young males (and students of both sexes), of adult native speakers when 'thinking privately' to themselves perhaps, and especially for the fantasy world of popular songs, foreign-movie subtitles, passionate letters, folktales, and the comics. This style, too (whether considered independently or in interaction with 'refined' speech) gives evidence that Malay address forms are rich with implications for the study of social hierarchy.

#### 4. THE ROLE OF CONSERVATISM IN THE RISE OF MODERN STANDARD MALAY

I shall henceforth take it as established that Malay is a special kind of T/V language with intimate ties to social hierarchy. My focus of interest in the next three sections is to present a perspective on the history of the Malay language that is consistent with this conclusion.

As mentioned in the Introduction, it has often been assumed that Malay is a simple and 'democratic' language, but little analysis has been offered to back up this kind of summary statement. In American scholarship, at least, B. Anderson's (1966) view, quoted earlier and repeated in part below, has been cited by linguists as well as non-linguists, hence seems to represent something like the standard view. In part, the standard view includes the idea that there were important linguistic factors behind the choice of Malay (over, say, Javanese) as the national language of Indonesia. Further, the standard view explicitly suggests that Bahasa Indonesia was developed out of a kind of pidgin into a modern political language; and that the choice was felicitous because "Malay had ipso facto an almost statusless character, like Esperanto, and was tied to no particular regional social structure. It had thus a free, almost 'democratic' character from the outset..." (Anderson 1966:104). Earlier I dealt with some sociolinguistic evidence against this interpretation of Malay's current status. In the remaining space I shall attempt to challenge the historical corollary, namely, the idea that at some time in the past Malay was nothing but a simple pidgin or Esperanto-like language.



In a strict technical sense, the idea that Malay was ever merely a pidgin or Esperanto is patently false because it implies a time when there were no native speakers. However, I do not think this extreme version of the standard view is actually believed by exponents of it. It is enough that the role of native speakers has been underemphasised in the story of Bahasa Indonesia's development, and that correspondingly the role of non-native speakers has been overemphasised. I will argue against these emphases, and suggest that what eventually came to be the recognised standard for Bahasa Indonesia must have been substantially influenced by models provided by educated Malays on the mainland and on Sumatra, some of whom (such as the poet Amir Hamzah) spoke 'High' Malay natively (see below). On the other hand, it is clear that native speaker 'experts' did not simply dictate the terms of the emerging national language. The prevailing sentiment was aptly expressed in a poem by Rustam Effendy, himself a Sumatran (native speaker of Minangkabau Malay). The following was written around 1925:

I am not a slave of this land  
Bound by the laws of the experts.  
I reject the rules of grammar  
The structure of the old poems.

Translation by A. Teeuw (1967:19)

Teeuw rightly interprets Effendy's poetry as a declaration of independence from external standards, presumably including the particular programme promoted by the Dutch through Van Ophuysen's Malay grammar based on 'High' Malay.

Experimentation with language does not imply throwing language to the wind, however. Effendy's choices were extremely limited, and revolved around a few longstanding issues. For more than a hundred years, the Dutch colonial government's need to solve the language problem included some halting efforts at standardising Malay (Hoffman 1979). As Teeuw (1967:7) puts it, eventually "Malay took its irreplaceable position as the vehicular language, even though this solution to the language problem was neither consciously intended nor officially chosen by the Dutch". It is thus no wonder that the issues were simply carried over into the early nationalist period.

As reported by Rafferty (1989:1), "radically opposing views" as to the origin of modern standard Indonesian have been proposed by Teeuw on the one hand, and the prominent Indonesian scholar, S.T. Alisjahbana on the other. Alisjahbana's opinion (1962:1) squares with the standard view, cited above: "In a short span of time", he writes, "this language (Indonesian) has been transformed from an un-integrated, pidgin-like lingua franca into an official language". In contrast, Teeuw has maintained that the basis of Indonesian was the language encouraged by Dutch scholars, and promoted under the literary umbrella of the Balai Pustaka publishing house (1920-1942). The style of Malay promoted by the latter, called Balai Pustaka Malay (BPM) by Rafferty, was based on 'High Malay'. According to Rafferty, 'High Malay' is a cover term that includes classic literary Malay as well as the modern dialects spoken on both sides of the Malacca Strait, in Riau, Lingga and Johor. Rafferty notes that BPM was "based on van Ophuysen's Malay grammar which became the standard for teaching Malay at the time" (Rafferty 1989:1).

It is of interest in this paper to point out the narrow focus of this debate. The focus is on 'High' Malay (spoken by native speakers in the area around the Malacca Strait) versus 'Low' Malay (spoken as a lingua franca by non-native speakers). Totally out of contention at that time, apparently, were the hundreds of regional varieties of Malay spoken by millions of native speakers. These included, among others, Palembang Malay, Bengkulu Malay, Minangkabau, and Jakarta Malay (Betawi).

These varieties have been given the collective name 'Middle Malay', but in no way do they resemble a coherent grouping. The question that immediately comes to mind is, why did the debate over a suitable standard centre only on 'High' versus 'Low' Malay?

One likely reason is that the so-called 'Middle Malay' dialects, by the turn of the century, displayed so much diversity – that is, had changed so drastically in phonology and morphology – that they were unintelligible to outsiders. (And so they remain today. See Errington (1986) for an illuminating recent study of contrasts between Indonesian and Jakartanese Malay.)

Interesting linguistic evidence has recently come to light which supports this suggestion, and which throws light on the possible role of linguistic conservatism in the development of modern Bahasa Indonesia. According to Blust (1981), modern Standard Malay is extremely conservative when compared with other languages of the Malayo-Polynesian family. This is a remarkable conclusion in light of the debate outlined above. It implies that the language that emerged (Indonesian) shows little sign of 'language mixing' in key areas of structure. If true, this fact would be difficult to reconcile with the standard view that Indonesian simply arose from totally unintegrated, pidginised varieties.

Blust's evidence is drawn from pronoun forms, phonology, and 'basic' vocabulary. 'Basic' vocabulary in this context involves everyday words like 'two', 'five', 'eye' and 'louse'. Common vocabulary items are generally assumed to be relatively stable and resistant to borrowing. Based on a fair sampling of Malayo-Polynesian languages in Southeast Asia and the Pacific, Blust found that Standard Malay has changed less in the course of its separate history than any of the 800 or so languages in this family. Not only does Standard Malay retain *kau*, *kamu*, and several other pronouns from Proto-Malayo-Polynesian, it displays a startling 59% retention per cent of basic vocabulary. This retention percentage is compared to 40% for Tagalog (Philippines) and 30% for Jogjakartan Javanese. The next closest rival was found to be Minangkabau (50%). Minangkabau is arguably a dialect of Malay, but it has undergone a number of structural changes that render it unintelligible to other Malays, hence highly 'marked' as a local variety.

Even allowing for a wide margin of error in Blust's study, the implication for recent history is driven home quite forcefully. While playing host to multitudes of second-language users for centuries, and after intensive 'development' efforts to transform it into modern Bahasa Indonesia, the resultant language continues to exhibit more relative stability than any other language or dialect of the entire Malayo-Polynesian family of languages. This fact suggests strongly that although the language has changed and will continue to change, as does any living language, Malay remains relatively conservative at the core. For this to be possible, some recognisable core must have been held in high esteem for a very long time; otherwise, the hosts of second-language users would surely have had a more drastic effect on basic vocabulary and structure.

If accepted, Blust's conclusions would have to be regarded as little short of miraculous by anyone who holds the standard view that modern Bahasa Indonesian developed out of a mere pidgin or trade language.

##### 5. REMARKS ON THE 'SUITABILITY' OF MALAY FOR ITS ROLE AS THE BASIS OF THE NATIONAL LANGUAGES OF INDONESIA AND MALAYSIA

In this section, I will develop a perspective on the history of Malay that is consistent with the idea, outlined in the previous section, that Standard Malay is relatively conservative as languages go.

The fundamental assumption I make is that the Malay heartland is the 'High' Malay region straddling both sides of the Malacca Strait. As it happens, for better or for worse as far as the Malays themselves are concerned, this region has always been of supreme international importance. Accordingly, many major historical 'centres' have been found in this region, for example at Palembang, Jambi, Johor, and Malacca, to name only a few (Coedes 1968).

The native language spoken on both sides of the Malacca Strait is of course Malay. I take it for granted that there is a certain naturalness to the suggestion that, historically, settlers who came to this region tended to be absorbed into the land and to adopt the Malay language and culture as their own. The pattern must have been repeated many, many times in the past. If so, then the term 'Malay' itself must clearly mean language and culture, not race. Likewise, historical mixing of races does not necessarily imply a hybrid language and culture. Only history can guide us here. Whether and how the Malay language and culture either shaped, or were shaped by, in-migrations, invasions, conquests, religious movements, and tourism is a question for research. Below I shall pose a few specific questions that future researchers might consider.

To account adequately for the position of modern Standard Malay, I believe a much greater explanatory role must be given to its continuity as an international language than has been accorded in the recent literature. (A notable exception is Teeuw 1967.) Scholars should take much more seriously the evidence of Malay's prestige in the archipelago. Most importantly, the temptation must be resisted to compare Malay's prestige with that of Dutch and English during the early colonial period. This period was, after all, one during which *all* peoples and cultures in the region underwent a decline of autonomy and loss of dignity. Accordingly, less importance should be given to the fact that Malay was pidginised and widely used as a trade language; and correspondingly, more weight should be given to the evidence that other languages of the area were never pidginised. In this context, it seems relevant to point out that only major international languages typically give rise to lingua francas in the first place. Latin, English, French, Spanish, Portugese, and Malay have given rise to such varieties. Why is it that German, Swedish, Japanese, Hindi, and Tagalog do not belong in this list? And, closer to the point, why has Javanese never been pidginised?

Once the continuity of Malay's historical role is considered seriously as a working hypothesis, it seems to matter much less whether Malay possesses (or lacks) this or that linguistic feature. In what follows, I will try to suggest the kinds of sociolinguistic generalisations which should be sought in future research.

Within the point of view I am developing, it seems important to underscore the fact that Malay's prestige, although seriously undermined by the Dutch presence during the colonial period, has apparently never been seriously challenged by another language in the region. In a significant sense, Malay's relative position has not been challenged since the seventh century.

Especially important are the bits and pieces of evidence concerning life in the archipelago between the seventh and fourteenth centuries. As pointed out by Lowenberg (1990:110):

The first institutionalized spread of Malay occurred during the Srivijaya Empire (seventh through fourteenth centuries A.D.) which adopted Malay as its official language. From its capitol at contemporary Palembang in southern Sumatra and a secondary base at Kedah on the Malay Peninsula, Srivijaya eventually conquered all of Sumatra, West and Central Java, and the Malay Peninsula, and established colonies along all seacoasts and major rivers within its domain. It maintained diplomatic relations with both India and

China and effectively controlled both the Straits of Malacca and the Straits of Sunda for over five centuries.

The only problem I have with the above passage is the phrase 'adopted Malay as its official language'. What does it mean to say that Srivijaya 'adopted' Malay? Why not say that Malay was the language of the Empire? The Romans did not 'adopt' Latin as the language of the Roman Empire; Latin was ipso facto the language of the empire, by virtue of the fact that the centre was in Rome. Likewise Malay, as the native language of the Srivijaya capital at Palembang, Sumatra, was ipso facto the language of power.

Lowenberg's choice of terms betrays a second false assumption that I believe to be both unnecessary and unsupported by any evidence. I will call this assumption the 'phoenix' theory and oppose it to the 'continuity' theory which I defend. Since Malay was the language of Srivijaya, everyone acknowledges that Malay must once have had great prestige in the region. However, according to the phoenix theory, Malay's prestige thereafter died without leaving as much as a trace (much as did memory of Srivijaya itself).<sup>23</sup> There are variants of the phoenix theory, of course, but some scholars seem to hold a very strong version of it. Its utility is that it sets the stage for a dramatic story of the 'creation' of modern Malay (out of its own ashes, so to speak) by visionaries and 'language engineers'. The phoenix theory underlies typical summary statements like the following.

...the monsoon pattern made it impossible to complete the voyage (between China and India) without a pause of some months in the Malay-speaking region, a fact which resulted in Malay eventually acquiring the status of lingua franca throughout the Archipelago. (Prentice 1987:911)

If one is to believe the above statement, one is forced to conclude that Malay would not be widespread were it not for the language-learning efforts of overwintering foreign sailors. Moreover, one is left to imagine for oneself how these same sailors would have been able to convince other non-Malays to use this language over a vast geographic area roughly the size of the United States.

The obvious objection lies in the likelihood that the author of this often-repeated idea has got the implication exactly backwards. The sailors did not play a causal role in the spread of Malay; they found it convenient to use Malay because Malay was already widespread. What would motivate them to learn the language otherwise? Prentice is apparently aware of a gap in the standard argument, since he qualifies the above with his next sentence:

Although this expansion of the language has not been documented, it is known that Malay was already in use in eastern Indonesia in the sixteenth century and it was considered quite normal for Francis Xavier to preach in Malay when he was in the Moluccas.

The reader should be mindful of the fact that the Moluccas are 2,000 miles from the Strait of Malacca. The author's mention of an undocumented "expansion of the language" betrays his acceptance of a rather strong version of the phoenix theory.

According to Lowenberg's (1990:111) account, quoting Alisjahbana (1976:33-34):

...in 1614 Jan Huygen van Linschoten, a Dutch navigator, observed that 'Malay was not only the most prestigious of the languages of the Orient...he who did not understand it

<sup>23</sup>The existence of Srivijaya was discovered in this century by Western scholars studying Chinese historical documents (Coedes 1968).

was somewhat in the same position as Dutchmen of the period who did not understand French.'

Surely such documents point to something important in terms of prior conditions. Why did Malay have such great prestige in 1614? It seems simplest to suppose that Malay's prestige had not been lost in the archipelago. Indeed, the available documents are perfectly consistent with the assumption that a rather direct line should be drawn between Srivijaya and modern Standard Malay.

Much evidence gathered by Prentice and Lowenberg can be used to argue against the phoenix theory. As they note, many Empires have come and gone in the region since the seventh century. But the geopolitical 'centre' of the archipelago – the lands and waters over which each successive Empire contested – has remained the land base from which to control the Strait of Malacca. The record lends no support to the idea that civilisation has ever declined in this region since the seventh century. There were many successive governments that followed in train, as is well known. Later empires originated as expansionist movements in Java; after Srivijaya (AD 650-1350) came Majapahit (AD 1293-1500), and Mataram (AD 898-1750). All three medieval empires had major centers in Sumatra and Malaya that competed with and sometimes launched bloody wars against Java (Coedes 1968:144). The last of the great empires to survive, Mataram, was brought to heel by the Dutch in the eighteenth century.

An interesting twist is of course the central role played by non-Malays in the Malay homeland, especially the Javanese. It need scarcely be questioned that Java's greatest export in the region has been government. For example, some ethnic groups of interior Sumatra even claim an affinity with medieval Javanese empires. The highland Rejangs of Bengkulu Province claim their culture to be linked directly with Majapahit. The Rejangs record in their oral history that a major change in their customary laws occurred when four 'princes' (*pangeran*) of Majapahit, after losing a dispute at court, fled to the hinterland in search of territory to rule. The four princes offered the people what they knew best: government. According to legend, the princes were freely elected by the Rejangs, who thereby overthrew their old customary laws based on the absolute power of the *ajai*'s (Hosein 1971). To complete the story it must be assumed that the princes originated in one of the 'centres' in Sumatra (such as Jambi), and spoke and were answered in Malay. Without these assumptions, the story could scarcely be considered possible.

Even if legends like the above are rejected, historical documents record a similar pattern in many other places in Sumatra and the Malay peninsula. According to Coedes (1968:245), the founder of the Sultanate of Malacca was a certain Parameśvara, a native of Palembang and the husband of a princess of Majapahit. (Parameśvara was probably ethnically Javanese himself, but he was at the same time 'a native of Palembang' as Coedes reports.) Failing to seize power in Palembang after the death of Hayam Wuruk (1389), he took refuge at Tumasik (Singapore) and killed its Malay chief, a vassal of the Siamese kingdom ruling there at the time. After reigning a few years, he was driven out of Singapore by the Siamese. He and his court fled first to Muar, then to Bertam, and finally to Malacca where a permanent kingdom was established.

The story of the Sultanate of Malacca is continued by Prentice (1987:916) as follows:

After the defeat of Malacca by the Portugese in 1511 the court fled to the south and eventually established a polity which embraced Johor and the island groups of Riau and Lingga in modern Indonesia.

Prentice goes on to explain the relationship of these political developments and the Malay language:



The literary traditions of the Malacca sultanate survived the upheavals of the colonial period and continued at the court of the Sultans of Riau-Johor...The literary Malay of the court continued...to be regarded as the standard on both sides of the frontier and served in both areas as the basis for the future national language.

While the historical role of the Javanese in the Malay heartland has no doubt always been of extreme importance in the region, one crucial assumption must be added to gain closure in all these cases. That is, it is necessary to assume that business in the Malay heartland was conducted in a commonly understood language at each particular point in time. I know of no other possible candidate than Malay. Moreover, many who came as conquerors remained as settlers. If they did not impose their own diverse languages on the region, the implication is that the region absorbed them. Like the Normans of England,<sup>24</sup> many Javanese came as invaders and were eventually absorbed into the land.

The historical record thus presents a picture of intensive non-Malay activity in the region – Javanese, Chinese, Arab, European, and recently Japanese and American – but no break (save during the colonial period) in the continuity of Malay's position as the language of the 'centre'. Importantly, to maintain this picture, it is not necessary to assume that the Malays' political role was comparable in importance to their linguistic and cultural contribution.

In 1824 the British and Dutch split the Malacca Sultanate into two, leaving one half on the Dutch side and the other on the British side. This action no doubt contributed further to the precipitous fall in prestige of all native languages, including Malay, in relation to the European languages. There is ample evidence of early colonialist attempts to further denigrate Malay. It is sad to consider the possibility that echos of this anti-Malay ideology may still linger on in the standard view of the history of Malay language and culture.

While there seems to be no truth to the phoenix theory applied to medieval times, it seems nevertheless to be true that 'native' prestige must have reached an all-time low after 1824. However, by 1900 it was clear that an anti-Malay campaign in Indonesia had failed. That the failure was assisted in part by Dutch progressives does not alter the point being made here (Hoffman 1979). Beginning in 1886 the Dutch, goaded by some of their own scholars and by liberal movements in Holland, began promoting the use of Malay actively in the colony, which of course helped to set the stage for nationalist revolution. It seems that no other language of the region has ever posed even a remote challenge to the position of Malay.<sup>25</sup>

## 6. IMPLICATIONS FOR LANGUAGE PLANNING

Issues that warrant further study concern general sociolinguistic patterns associated with languages that function in inter-ethnic communication. One fairly recent line of research, called Accommodation Theory, has been launched by Giles (1979). Accommodation Theory investigates (among other things) the ways in which diverse ethnic groups behave when forced by circumstances to speak a common second language. The term 'Accommodation' here refers to the fact that intergroup and inter-ethnic relationships are subject to constant negotiation during the course of each interaction (Ellis

<sup>24</sup>The same relationship probably holds of Mandarin Chinese in relation to the long succession of rulers in Peking.

<sup>25</sup>Including, of course, Javanese. The paradox of Javanese is of course that whereas the culture has produced several great empires, the language has never been widely adopted by non-Javanese. This fact complements the Malay paradox studied in this paper. It is unlikely that either paradox can be resolved independently of the other.



1986:256). Below I suggest five assumptions that might be explored from the general perspective of non-Malays using Malay as a second language, and one from the perspective of native speakers.

First, a high level of prestige must be assumed to be attached to knowledge of Malay in the minds of non-Malays. This prestige must have existed for centuries, even when the numbers of proficient non-native speakers of Malay reached its lowest ebb, as it presumably did during the early colonial period. This factor would serve to motivate non-Malays to use the language for inter-ethnic communication; and some to try to master it to the best of their ability. Second, the practical utility of Malay as a lingua franca has probably always meant that to speak Malay even badly is nonetheless advantageous – another clear sign of Malay's prestige. Thirdly, the centres of Malay language and culture, while shifting from time to time amongst numerous sites in Sumatra and peninsular Malaya, have nonetheless always been recognisable and available to provide a model and standard of 'correct' style in both speech and writing. Fourth, there is plenty of evidence that Standard Malay has changed in many ways over the centuries. Based on its long history of writing, which can be dated from AD 683,<sup>26</sup> it is clear that Malay has undergone many linguistic and cultural changes internal to itself. Finally, Malay dialects relatively distant from the current centres of power have diverged even further, and have become thereby highly 'marked' as having merely local currency.<sup>27</sup>

Another research topic comes to mind that takes the perspective of native speakers. What characterises native speakers of widely-used languages? What are their attitudes towards the fact that their language is used as the medium of inter-ethnic communication? One would not be surprised to find that most native speakers of Standard Malay are habituated to the fact that their native language has inter-regional and inter-national 'responsibilities' and status. This fact seems automatically to induce tolerance of second-language users (Giles 1979). While the psychological traits of openness and tolerance are considered universals in Accommodation Theory, their enhancement in native Malay speakers could be investigated and verified as another noteworthy contribution of Malay culture in Southeast Asia.

## 7. CONCLUSION

The political task of this century was to build a new Indonesian nation the size and scope of Srivijaya and Majapahit. The task called for a unifying language. To select a language, and to determine what form it should take, politicians, poets, and language planners, while wrangling over details, re-discovered an ancient channel provided by history.

If the arguments of sections 1-3 of this paper are accepted, then some of the questions raised in sections 4-6 may be clarified. In particular, it should no longer be regarded as problematic to suggest that the system of address forms in modern Bahasa Indonesia carries strong implications for the study of hierarchy in social structure.

---

<sup>26</sup>Prentice (1987:915).

<sup>27</sup>For example, Palembang-Malay is quite divergent from Standard Malay. My comparison using a modified Swadesh 200 item word list indicates only 80% shared 'basic' vocabulary between Palembang-Malay and Bahasa Indonesia (=Standard Malay as it is recognised in Indonesia). This implies that Palembang, although as the seat of the Srivijaya Empire it represented the 'centre' from the seventh to twelfth centuries, is today relatively distant from the *current* centre of Malay language and culture. This implication seems consistent with the facts.

APPENDIX: SECOND PERSON PRONOUNS IN SELECTED INDONESIAN LANGUAGES

In this appendix are listed the second person pronouns in common use in a number of Indonesian languages. See Appendix 3 for sources of data on these languages.

1. NON-MALAY

Language	Province	1SG	2SG	2PL	Honorific
Acehnese	Aceh	<i>kee</i>	<i>kah</i>	<i>kah</i>	<i>droe</i> <sup>28</sup>
Batak	North Sumatra	<i>ahu</i>	<i>ho</i>	<i>hamu</i>	<i>hamu</i>
Rejang	Bengkulu	<i>uku</i>	<i>ko</i>	<i>udi</i>	<i>kumu</i>
Lampung	Lampung	<i>nya'</i>	<i>niku</i>	<i>kuti</i>	<i>pusi-kam</i>
Kawi (Old Javanese)		<i>aku</i>	<i>ka(N)u</i>	<i>ka(N)u</i>	<i>kita</i>
Tengger-Javanese	East Java	<i>aku</i>	<i>sira</i>	<i>sira</i>	<i>rika</i>
		<i>i(ng)sun</i>			
		<i>(r)éyang</i>			

2.(a) 'CLASSIC' T/V DIALECTS OF MALAY (KAMU - HONORIFIC)

Language	Province	1SG	2SG	2PL	Honorific
Besemah	South Sumatra	<i>aku</i>	<i>kau</i>	<i>kamu</i>	<i>kamu</i>
Benakat	South Sumatra	<i>aku</i>	<i>kau</i>	<i>kamu</i>	<i>kamu</i>
OKI	South Sumatra	<i>aku</i>	<i>kau</i>	<i>kamu</i>	<i>kamu</i>
Belitung	Belitung I.	<i>aku</i>	<i>kau</i>	<i>kamu</i>	<i>ikam</i>

(b) MALAY DIALECTS THAT USE KAMU OR AN EQUIVALENT NON-HONORIFICALLY

Language	Province	1SG	2SG (masc)	2SG (fem)	2PL
Standard Malay	East Sumatra <sup>29</sup>	<i>aku</i>	<i>(ENG)KA&lt;-M&gt;U</i>	<i>(ENG)KA&lt;-M&gt;U</i>	<i>kalian</i>
Bengkulu	Bengkulu	<i>aku</i>	<i>kamu</i>	<i>kau</i>	<i>kamorang</i>

Language	Province	1SG	2SG (masc)	2SG (fem)	2PL
Ujan Mas	South Sumatra	<i>aku</i>	<i>kaba/denga</i> <sup>30</sup>	<i>kaba/denga</i>	<i>kamu</i>
Palembang	South Sumatra	<i>tubu</i>	<i>awak</i>	<i>awak</i>	<i>kamu</i>
Ogan Ulu	South Sumatra	<i>aku</i>	<i>ngan</i>	<i>ngan</i>	<i>ngan</i>
Kayo Aro	South Sumatra	<i>aku</i>	<i>nga</i>	<i>nga</i>	<i>nga</i>
Jakarta	Jakarta	<i>gue</i>	<i>kamu</i>	<i>kamu</i>	<i>sekalian</i>

<sup>28</sup>Achenese and Lampung express honorifics in all the pronouns, not just second person.

<sup>29</sup>The three northernmost Malay-speaking provinces of Indonesia lie along the Strait of Malacca and are said to use a variety of Malay natively that most resembles the 'received' Standard. They are the provinces of North Sumatra, Riau and Jambi.

<sup>30</sup>*kaba* is used to refer to same-sex individuals; *denga* is used for opposite-sex individuals.

## 3. SOURCES OF DATA FOR THIS APPENDIX

## (a) MALAY DIALECTS (FW = fieldwork conducted by the author)

Language/Dialect	Province	Principal Source
Bengkulu	Bengkulu	FW, Amran Halim
Benakat	South Sumatra	Nangsari Achmad <sup>31</sup>
Besemah	South Sumatra	FW, Gaffar (1983)
Belitung	Belitung I.	FW, Husadi Fitoy
Kayo Aro	South Sumatra	FW, Chuzaimah Diem
Ogan Komerang		
Iilir (OKI)	South Sumatra	FW, Moh. Junus
Ogan Ulu	South Sumatra	FW, Neli
Palembang	South Sumatra	FW, Amir Faizal
Ujan Mas	South Sumatra	FW, Nangsari Achmad

## (b) NON-MALAY LANGUAGES

Language/Dialect	Province	Principal Source
Acehnese	Special province of Aceh	Durie 1985
Batak	North Sumatra	Van Der Tuuk 1971
Rejang	Bengkulu	McGinn 1982
Lampung	Lampung	Walker 1976
Standard Javanese	Central, East Java	Dardjowidjojo 1978
Tengger-Javanese	Special territory of East Java	Smith-Hefner 1988

## BIBLIOGRAPHY

- ALISJAHBANA, S. T., 1976, *Language planning and modernization: the case of Indonesian and Malaysian*. The Hague: Mouton.
- 1978, The concept of language standardization and its application to the Indonesian language. In Perez, Santiago and Nguyen Dang Liem, eds, 1978:19-42.
- ANDERSON, Benedict R., 1966, The languages of Indonesian politics. *Indonesia* 1/1:89-106.
- ANDERSON, Edmond A., 1983, The meaning of variation in Indonesian. *NUSA* 15:1-26.
- BECKER, A.L. and I Gusti NGURAH OKA, 1974, Person in Kawi: exploration of an elementary semantic dimension. *Oceanic Linguistics* 13/1-2:229-255.
- BELLWOOD, Peter, 1985, *Prehistory of the Indo-Malaysian archipelago*. Orlando, Fla: Academic Press.
- BLUST, Robert A., 1981, Retention rates in Austronesian languages. MS. (Cited in Bellwood 1985:118).
- BROWN, Roger and Albert GILMAN, 1960, The pronouns of power and solidarity. In Thomas A. Sebeok, ed. *Style in language*, 253-276. New York: Wiley; also in Pier Paolo Giglioli, ed. *Language and social context*, Harmondsworth, Penguin, 1972, 1976.
- COEDES, G., 1968, *The Indianized states of Southeast Asia*. Honolulu: The University of Hawaii Press.
- DARDJOWIDJOJO, Soenjono, 1978, Javanese honorifics and their linguistic implications. In the author's book of essays, *Some aspects of Indonesian and Javanese linguistics*. Honolulu: Department of Indo-Pacific Languages.

<sup>31</sup>Dr Nangsari is not a native speaker of Benakat-Malay, but a linguist who reported the data to me.

- DURIE, Mark, 1985, *A grammar of Achenese on the basis of a dialect of North Aceh*. Dordrecht: Foris.
- ELLIS, Rod, 1986, *Understanding second language acquisition*. Oxford: Oxford University Press.
- ERRINGTON, J. Joseph, 1986, Continuity and change in Indonesian language development. *Journal of Asian Studies* 45/2:329-353.
- FREIDUS, Alberta Joy, 1977, *Sumatran contributions to the development of Indonesian literature, 1920-1942*. Honolulu: University of Hawaii Press.
- GAFFAR, Zainal Abidin, 1983, *Morfologi dan sintaksis bahasa Besemah*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- GEERTZ, Clifford, 1960, *The religion of Java*. Chicago: University of Chicago Press.
- GILES, H., 1979, Ethnicity markers in speech. In Klaus R. Sherer and Howard Giles, eds *Social markers in speech*. Cambridge: Cambridge University Press.
- HALIM, Amran, 1978, The vernacular languages in relation to the standardization of Bahasa Indonesia. In Perez, Santiago and Nguyen Dang Liem, eds, 1978:297-303.
- HOFFMAN, J. E., 1979, A foreign investment: Indies Malay to 1901. *Indonesia* 27:65-92.
- HOSEIN, H. M., 1971, Rejang asal-usul. 58pp mimeo, edited by Abdullah Sani.
- JENSON, Kenneth M., 1988, Forms of address in Indonesian. *Review of Applied Linguistics* 81-82:113-138.
- KASWANTI PURWO, Bambang, 1984, The categorial system in contemporary Indonesian: Pronouns. *NUSA* 19:55-74
- KRIDALAKSANA, Harimurti, 1974, Second participant in Indonesian address. *Language Sciences* 31:17-20.
- LOWENBERG, Peter H., 1990, Language and identity in the Asian State: the case of Indonesia. *The Georgetown Journal of Linguistics* 1/1:109-120.
- McGINN, Richard, 1982, *Outline of Rejang syntax*. Jakarta: NUSA 14.
- McGINN, Richard, ed., 1988, *Studies in Austronesian linguistics*. Ohio University Monographs in International Studies 76. Athens: Ohio.
- PEREZ, A.Q., A.O. SANTIAGO and NGUYEN DANG LIEM, eds, 1978, *Papers from the Conference on the standardization of Asian languages, Manila, Philippines, December 16-21, 1974*. Canberra: PL, C-47.
- PRENTICE, J. D., 1987, Malay Indonesian and Malaysian. In Bernard Comrie, ed., *The world's major languages*, 913-935. New York: Oxford University Press London: Croom Helm.
- RAFFERTY, Ellen, 1989, The formation of the Indonesian national language in the early twentieth century. (To appear in *NUSA*.)
- SMITH-HEFNER, Nancy, 1988, Cara Tengger: notes on a nonstandard dialect of Javanese. In McGinn, ed. 1988:203-234.
- TEEUW, A., 1967, *Modern Indonesian literature*. The Hague: Martinus Nijhoff.
- 1972, The impact of Balai Pustaka on modern Indonesian literature. *Bulletin of the School of Oriental and African Studies* 35/1:111-127.
- TUUK, H. N. van der, 1971, *A grammar of Toba Batak*, translated by Jeune Scott-Kemball. The Hague: Nijhoff.
- WALKER, Dale F., 1976, *A grammar of the Lampung language: the Pesisir dialect of Way Lima*. Jakarta: NUSA 2.

The first part of the book is devoted to a general history of the United States from its discovery by Columbus in 1492 to the present time. It covers the early years of settlement, the struggle for independence, the formation of the Constitution, and the development of the nation as a great power. The second part of the book is devoted to a detailed history of the United States from 1789 to the present time. It covers the early years of the Republic, the struggle for reform, the Civil War, and the Reconstruction period. The third part of the book is devoted to a detailed history of the United States from 1865 to the present time. It covers the Reconstruction period, the Gilded Age, the Progressive Era, and the modern era.

The book is written in a clear and concise style, and is suitable for use in schools and colleges. It is a valuable source of information for anyone interested in the history of the United States.

## SIXTEEN NUCLEAR MICRONESIAN VERBS

JEFF MARCK

### 1. INTRODUCTION

Sixteen verbs are compared through nine Nuclear Micronesian languages for the development of idiomatic phrases centring on the causative prefix and directional suffixes. Findings are compared to those of Jackson (1983) for subgrouping Nuclear Micronesian. A settlement model for the area is proposed suggesting language differentiation along what Pawley and Green (1984) would call a radiation model overlaid by a period of areal influences in eastern Micronesia that appears to have fallen short of what Pawley and Green (1984) would call a network-breaking situation. The method of idiomatic phrase comparison is evaluated.

### 2. BACKGROUND

As a problem in culture history, the dispersal of Austronesian-speaking people out of Insular Southeast Asia into the Admiralties, the Bismarck Archipelago, over the Solomons and Vanuatu and out to New Caledonia, Fiji, Polynesia and parts of Micronesia is an object of fascination. Monolingual societies consolidated by royalty in Polynesia and parts of Micronesia are closely related in the archaeological record to fragmented linguistic and political landscapes in Melanesia and other parts of Micronesia (cf. Bellwood 1978, 1984; Green 1976, 1981; Green and Mitchell 1983; Irwin 1980; Kennedy 1982, 1983; Pawley 1981, 1982; Pawley and Green 1973, 1984; Shutler and Marck 1975).

We are confronted with a situation where the early parts of this dispersal could have been accomplished, theoretically, with relative ease due to the geographical configuration found between the Bismarck Archipelago and the south-east Solomons. Land is visible continuously along most of this stretch and it could have been settled as soon as the demographics of the early communities allowed for recruitment of people to push deeper into the Pacific. As a problem in linguistic reconstruction, it seems that the movement of people across the area and down the lines of dispersal occurred fairly rapidly. One need only look at an Oceanic family tree (cf. Grace 1955; Dyen 1965; Pawley and Green 1984) to understand that lifetimes of work by competent people have yet to dent the subgrouping problem in terms of the highest order of relationships.

In the case of Micronesia, we have the dual quandary of determining the internal relationships of the nuclear group and of identifying its nearest relative in Melanesia. An apparent early internal



diversification of Nuclear Micronesian out of a dialect or language in Melanesia that was not markedly different from other Oceanic languages as they were spoken at the time appears likely. Nuclear Micronesian shows no broad set of distinctive characteristics that allow its subgrouping with other Oceanic languages (cf. Blust 1984) in a manner comparable to Polynesian with some Fijian languages or the grouping of these with some Vanuatu languages. It is likely, therefore, that the discovery of Micronesia and the isolation of languages there was part of a kind of demographic explosion that occurred rather 'instantaneously' (Irwin 1980:326) by no later than 1600 BC out of the Oceanic heartland area.

Pawley (1981, 1982) has shown that there were elevated rank concepts in these communities, relative to what we find in Melanesia today. One can imagine that such institutions may have consolidated the settlement process and effected the development of local prestige dialects and the ultimate linguistic configuration of the islands as their population densities rose through natural growth and continued immigration.

Micronesia appears to have been part of this mid second millenium BC population explosion into the remote island environment. At the time of initial European contact, it had monolingual societies with royalty in the eastern islands of Kosrae, the Marshalls and Kiribati (the area of Nuclear Micronesia where differentiation among neighbouring languages appears to be oldest). Certainly the lack of historically optimal environments (abundant reef/lagoon and river estuary environments, cf. Bellwood 1978; Pawley and Green 1973, 1984) on the atolls (in the Marshalls, Kiribati and the Western Carolines) and their dearth even on the volcanic islands (Kosrae, Ponape and Truk) necessitated major subsistence adjustments in the earliest community and created pressure to spread to new areas as soon as they were known. The archaeology of Micronesia suggests fairly complete settlement of the area by the beginning of the first millenium AD but these researches remain in their infancy (Jackson 1983; Shutler pers. comm.). A language pattern suggesting second millenium BC settlements is not inconsistent with the preliminary nature of the archaeological work.

Blust's (1984) comparison of Nuclear Micronesian to Cristobal-Malaitan suggests only the faintest level of special relationship, and a minimal period of common development before Nuclear Micronesian became distinct from a potential common ancestor. Jackson's (1983) trenchant analysis of Nuclear Micronesian languages, as presently described, produced only moderate evidence of a unified Oceanic subgroup as compared, for instance, to the abundant innovations that are distinctly Polynesian. Clearly, things were happening quickly as Proto Micronesian emerged out of other Oceanic languages and then began its internal diversification.

### 3. METHODOLOGY: PHRASE LEVEL COMPARISONS

Jackson (1983) discusses what we mean by 'Nuclear Micronesian' and how the term developed in the literature. It is a subgroup of Oceanic Austronesian that includes all the languages of Micronesia except Chamorro in the Marianas, Yapese, Palauan and the Polynesian outlier languages of Nukuoro and Kapingamarangi. Nauruan is not well described, but preliminary work suggests that it is one of the nuclear group (Nathan 1973, n.d.; Jackson 1983). It is not considered here as no native speakers were available for the study. The purported proto-language is simply known as Proto Micronesian.

In 1978 those of us working on Nuclear Micronesian already knew that the lexicons were not heavily marked with shared innovations, except in the case of a subgroup consisting of Ponapeic and Trukic. Developments in the phonologies appeared to be influenced by processes in motion in the

proto-language, whilst developments in the grammars appeared to be clouded by the natural logic of syntactic change. Therefore, I set out to produce something other than a standard phonological, lexical or grammatical comparison. In his teachings and his work, Grace (cf. especially Grace 1981) has admonished us that languages have, through time, been spoken by real people in real places under real and varying ecological, technological, social and political circumstances and that, far from existing to entertain linguists with intellectual puzzles, which are sometimes artefacts of our method, they exist for the purpose of saying things.

At the time I began this work, Biggs (pers. comm.) had expressed his fascination with certain agreements that were turning up in his study of Eastern Polynesian languages. They involved similarities and identity between little phrases, little ways of saying things, that occurred within Eastern Polynesian subgroups already established on the basis of other criteria. Taken by themselves, in the individual languages, they did not stand out because the semantics did not seem extremely peculiar. He hesitated to call them idioms because their divergence from the expected semantics of the phrases was not so very pronounced. They are, at any rate, things that Grace (1981:19) would say are understood "holistically...the way of knowing a linguistic form immediately and as a unit".

As I approached this work in 1978 I was in Hawaii and had the opportunity to work with native speakers of Marshallese, Kosraean, Ponapean, Lagoon Trukese, Hall Island Trukic, Satawalese, Woleaian and Saipan Carolinian.<sup>1</sup> Linguists familiar with many of these languages were also present at the time. I have recently had the assistance of speakers of Kiribati and have been able to add it to the study as well.<sup>2</sup>

I had been influenced by Starosta who was interested in models of verbs in the grammar. And I recalled Lee's (1974) treatment of directional suffixes in Kosraean and Ponapean where he found a striking similarity in what he called a 'psychological' extension of the suffixes to non-motion verbs. They rendered aspectual and other notions and this same extension is made to verbs of motion to impart the aspectual sense. I wondered if these occasionally produced the kinds of twists of meaning Biggs was finding in Eastern Polynesian and, under the supervision of Starosta, I set out to systematically study the application of the directionals to sixteen Nuclear Micronesian verbs. In addition I looked at meanings rendered by the causative. These comparisons produced a number of phrases with odd twists in the semantics, and several others emerged incidental to the constructions I was studying systematically. I wouldn't call them all idioms, but in each there is something stable in small extensions of the expected meaning that appears to have remained in more than one language over a long period of time.

I will now present the findings and follow with a discussion of their distributions in relation to Jackson's (1983) evidence for subgrouping these languages. I will then consider historical developments that would be consistent with Jackson's evidence and that developed here. I will conclude with a consideration of what these little phrases seem to offer in relation to other distributions and their apparent durability over time.

---

<sup>1</sup>I wish to thank the following native speakers and other consultants: Marshallese - Titus Arelong, Daniel John, Alee Term, Chris Loeak, Byron Bender; Kosraean - Elmer Asher, Truman Wakuk, Alister Tolenoa; Ponapean - Ewalt Joseph, Hiroshi John, Welber Santos, Ken Reh; Lagoon Trukese - Kimeuo Kimiuo; Hall Island Trukic - Tony Otou; Satawalese - Joe Tiucheimal; Saipan Carolinian - Rosa Warakai; Woleaian - Camillus Taiogmai.

<sup>2</sup>I would like to thank Tereb'ata Groves and Gordon Groves for their assistance with Kiribati.

## 4. DATA AND DISTRIBUTIONS OF INTEREST TO THE SUBGROUPING PROBLEM

## 4.1 INTRODUCTION

The sixteen verbs considered were:

PMC <sup>3</sup>		
<i>*rongo</i>		to hear
<i>*lako</i>		to go
<i>*tisa(ng)</i>		to press
<i>*ka-ine-ni-eti</i>		to aim something at something
<i>*fili</i>		to choose
<i>*fanga</i>		to give
<i>*inu(m)</i>		to drink
<i>*kani</i>		to eat
<i>*suusuu</i>		to bathe
<i>*añum-ia/lima(k)</i>		to bail
<i>*t'uu</i>		to meet
<i>*lapa</i>		to be big
<i>*maturu</i>		to be asleep
<i>*mate</i>		to be dead
<i>*mauru</i>		to be alive
<i>*matau</i>		to be scared

Lee's (1974) study of the directionals in Kosraean and Ponapean produced its most striking similarities in the extended or psychological uses of *\*la* 'away', *\*sake* 'up' and *\*sio* 'down'. I found that the extended meanings occur in an almost identical manner in much of Trukic, and that using Satawalese as an exemplary Trukic language at least the following sets of identical function can be listed for Kosraean, Ponapean and Trukic:

<i>*la</i> 'away'	depletion	extinction
	completion	surface (all)
	totalisation	
	out of consciousness	
<i>*sake</i> 'up'	activity	cooked
	initiated	inchoative
<i>*sio</i> 'down'	surface (dot)	

The reader may wish to refer to Lee if the sense rendered by these combinations with non-motion verbs is not obvious. It is a detailed discussion and space considerations do not allow its repetition here. Similar reasons motivate reference to Jackson (1983) for a consideration of forms the directional suffixes have taken in Nuclear Micronesian.

<sup>3</sup>Abbreviations used for languages and proto-languages are as follows: CAR – Saipan Carolinian, HAL – Hall Island Trukic, KIR – Kiribati, KSR – Kosraean, MRS – Marshallese, NMC – Nuclear Micronesian, PMC – Proto Micronesian, POC – Proto Oceanic, PON – Ponapean, PTK – Proto Trukic, SAT – Satawalese, TRK – Lagoon Trukese, TRKic – Trukic, WOL – Woleaian. Other abbreviations are: CAUS – causative, DEF – definite, DUR – durative.

Marshallese showed some evidence of the extended system but much less than Trukic. Kiribati showed it not at all. The Marshallese agreements with Kosraean, Ponapean and Trukic were:

- \**la* 'away' a completive sense is seen in a few fossilised forms. Its current non-directional use renders a meaning of 'to hurry at something'
- \**sake* 'up' activity initiated  
inchoative

One might question the extent to which native speakers were clear about the extended meanings of these forms but, indeed, all were quite clear about them. Such meanings either did not exist (as in Kiribati), existed widely and systematically (as in Kosraean, Ponapean and most of Trukic) or existed to a degree but only with certain directionals (as in Marshallese). Thus it seems that if this was not a characteristic of Proto Micronesian, it developed early enough in an interstage language to pass on to a number of the contemporary languages, either as a part of their common ancestry or, perhaps, as an areal phenomenon that emerged before those that show it today had diverged into markedly distinct languages.

In addition to agreements previously reported by Jackson (1983), three agreements are reported here that are clearly Proto Micronesian, eight are distributed through everything but Kiribati, one was found in MAR/PON/TRKic, four were identified through PON/TRKic and eight were found exclusively in Trukic.

#### 4.2 DISTRIBUTIONS OF INTEREST FOR CHARACTERISING VERBS IN PROTO MICRONESIAN

Very striking is a sort of fossilisation in the verbs for pointing, orientation and aiming. Because they are so clearly multimorphemic all the languages seem to agree as to which morphemes were used in Proto Micronesian. The evidence suggests:

\**in(e, V)-ni-eti*

as the verb for 'oriented towards, pointing towards, in line with' and the causative:

\**ka-in(e, V)-ni-eti*

as the verb for 'to aim something at something/someone'.

Proto Oceanic \**piliq* 'to choose, select' is currently found only in Ponapeic and Trukic where it retains the POC meaning. In these languages \**fili-sake* is the common minimal usage and this may have been characteristic of PMC.

The noun/intransitive verb/transitive verb paradigm for 'bailer, to bail, to bail something' is quite fascinating as it involves two forms that are widely reflected through NMC and are known but extremely obscure in other Oceanic Austronesian languages. Consider:

\**lima* n., a bailer

MRS	<i>lem</i>	but:	KIR	<i>anim</i>
PON	<i>liim</i>		KSR	<i>ahnom</i>
PTK	* <i>lima</i>			

\**lima* Vi, to bail  
 PTK \**lima* but: MRS *yanyen*  
 KSR *enyen*

(KIR and PON: no data)

\**lima(k)* Vt, to bail something  
 PON *limak* but: KIR *anima*  
 PTK \**lima-i* KSR *inihm*  
 MRS *yalim*

Geraghty (pers. comm.) has noted the relation of some of these forms to a Proto Austronesian form rarely reflected in Oceanic which would have had the form POC \**lima(k)*, while Blust (pers. comm.) notes the similarity of the others to an Admiralties form that would be reconstructed as \**añum-ia* for POC. Certainly we are talking about the same words in both cases. Given Jackson's (1983) subgrouping, the PMC form would appear to have been \**añum-*.

It is not unusual for a noun to pair with an intransitive verb in these languages and to contrast with a suppletive form for the transitive verb. Likewise, a common pattern is for the noun to have a suppletive form versus the transitive and intransitive forms of the verb. There is a grading, then, where KIR and KSR tend to use reflexes of something like Blust's \**añum-ia* for everything, PON and Trukic tend to use reflexes of Geraghty's \**lima(k)* for everything and MRS is intermediate. Such a state of affairs is not entirely surprising. In other instances MRS leans towards PON and Trukic when KIR and KSR do not, as Jackson has shown and as will be seen below.

#### 4.3 DISTRIBUTIONS THROUGH MRS/KSR/PON/PTK

PMC \**rongo* 'to hear' appears to be drifting around the edges of a regular verb of motion in the languages other than KIR. That is, it can combine in the other languages with the regular motion-directional suffixes to indicate the direction from which news or information is coming. The difference is that the directional indicates source rather than terminus.

PMC \**maturu* 'to sleep' may have had a lexicalised causative aside from the predictable 'to put someone to sleep' meaning. This second meaning is found in all but KIR and Lagoon Trukese. The idiomatic use refers to sneaking up on sleeping fish, birds or animals (to catch them). This is quite striking because the causative is used to speak of taking advantage of a pre-existing condition rather than causing that condition.

Similarly, \**ka-maturu papine* (lit: 'cause - sleep woman') appears to have meant for a man to sneak up on a sleeping woman, pretend to be her husband or lover, and have sex with her without awakening her so fully as to make her realise it was not her normal paramour in bed with her. There was a moment when several native speakers departed from a session in mad hysterics when they came to realise that this usage was not limited to their own language. Surely we can feel a measure of satisfaction as linguists that it was us, rather than other culture historians, who first described this aspect of early Micronesian social behaviour.

\**maturu* also has an interesting application of the extended directionals in all the languages but KIR. Probably because of the incompatibility of the semantics of the form with \**sake* (i.e. 'sleep-up') for the inceptive, meaning 'begin to sleep, fall asleep', the inceptive is derived with \**la*/\**lako*

and the expected 'completely' meaning is carried in the contemporary languages by such adverbials as 'heavy'.

PMC *\*mate* 'dead, diminished physiological state' is disambiguated in all the languages but KIR by using *\*mate-la/lako* to carry the 'to be dead' meaning. In KIR that form has the meaning of 'hopelessly adrift at sea, drifted away at sea and lost'. Note that the KIR form implies motion and the absence of a cognate form for KIR in this and other comparisons involving directionals may be systematic. That is, they may have once existed in KIR but were shed along with the systematic application of directionals with extended meanings.

PMC *\*mauru* 'alive, to be in an enhanced physiological state' occurred in a common idiom for mentioning the successful delivery of a child in PMC or some early interstage. *\*mauru-la/lako* is the form in KSR, PON and PTK while *\*mauru* with the causative renders this widely understood meaning in MRS. *\*mauru* is replaced by the innovative *\*manawa* in all of Trukic but the idiom remains the same. KIR does not appear to have such a meaning for the free form, the causative or the verb with the *\*la/lako* directional.

*\*mauru* with *\*sake* has two senses in all the languages other than KIR. The first meaning is 'to revive, come back to life', the second a common usage referring to an erect penis.

In the languages other than KIR the causative of *\*mauru* 'to be alive' has two senses disambiguated by *\*la* or *\*lako*:

*\*ka-mauru* to sustain someone (take responsibility for their sustenance, feed them, house them, represent them in the community, etc.)

*\*ka-mauru-la/lako* to save someone (from death, a fight, etc.)

The contrast is, perhaps, predictable, the *\*la/\*lako* signalling a single or completed event as opposed to a continuing action or relationship. It is the innovative form *\*manawa* which is found in Trukic, but again, the usages are the same.

#### 4.4 DISTRIBUTIONS THROUGH MRS/PON/PTK

Comparison of the languages indicates that PMC *\*lapa* 'to be big' had the particular meaning 'to be huge' when reduplicated in some interstage. This is true of MRS, PON and PTK. In KSR the form occurs only as an adverbial, and KIR speakers do not recognise this meaning for the reduplicated form.

#### 4.5 DISTRIBUTIONS THROUGH PON/PTK

There is a derivational prefix *\*li-* reflected in PON and PTK meaning 'a person who characteristically engages in' that is attached to verbs to derive adjectives and nouns. The *\*li-* prefix has similar usages in other Nuclear Micronesian languages and might reflect an old feminine marker (Bender pers. comm.) but the particular usage mentioned above appears limited to PON and Trukic. Additionally, there is an idiomatic deployment of the prefix with the reduplicated form of *\*fili* 'choose, select' used in negative constructions as the common idiom for 'it doesn't matter (of alternatives)':



PTK \*E tai li-fili-fili  
 3SG NEG li-choose-DUR  
 It doesn't make any difference.

PON Soot li-pili-pil  
 NEG li-choose-DUR  
 It doesn't make any difference.

PMC *\*mate* 'to die' occurs as *\*mate-sio* 'die-down' when referring to multiple deaths (as in a plague) while MRS and KSR use *\*mate la/lako* 'die away' for the same meaning and KIR constructs such a meaning in neither manner.

The form *\*mate mwuri ni* means 'to lust after' (lit. 'to die after of') while KSR uses *\*mate lako* and MRS uses *\*mate* with *kahkey* 'die about, because of' to render similar meanings.

Two forms concerning the political system were encountered and appear to be uniquely shared:

*\*ka-rongo-rongo* (lit. 'cause hearing', reflected as a noun meaning 'a public hearing' in PON and all of Trukic)

*\*ka-pwungu* (lit. 'cause straight/true', reflected as a noun meaning 'a court hearing' in PON and all of Trukic)

PMC *\*fanga* 'to give' means 'to confess' when combined with *\*sake* 'up (directional suffix)' in PON and PTK and this appears unique in the nuclear group.

#### 4.6 DISTRIBUTIONS WITHIN TRUKIC

PMC *\*rongo* appears to have developed an idiomatic use of the causative in PTK: *\*ka-rongo-rongo* which meant 'to make someone overhear something (i.e. something which you do not wish to tell them directly)'.

PMC *\*lako* 'to go' had a lexicalised causative PTK *\*ka-lako* which meant 'to release' rather than the 'to drive off, force off' meaning suggested for PMC by the other languages.

PMC *\*kani* 'to eat' had a lexicalised causative in PTK *\*ka-kangi* 'too entice fish to eat by feeding them treated bait'.

The *\*t'uu* 'to meet' form had a causative PTK *\*ka-t'uu* which meant 'to keep one's eyes open for' which contrasts with the expected meaning of that word in the causative 'to arrange for one party to meet another', a meaning current in the cognate PON form and possibly, therefore, in PMC.

PMC *\*lapa* 'to be big' occurs as an adverbial suffix meaning 'to much' in the sense of excessively or contrarily, and this meaning is not found other than in Trukic.

PMC *\*lapa* had a homophonous form in PTK: a verb of motion referring to the motions of coitus in all of Trukic represented in this study. It has special meanings with the directionals having to do with the direction of the motions or orientation of the actants, and when reduplicated refers to a woman's ultimate ecstasy (and also that of the man, apparently. Conversations surrounding this form had a tendency to break down quite readily). Native speakers of non-Trukic recognise none of this as part of their language, while all the Trukic speakers instantly produced a number of variations of the form.

PMC *\*mate* 'to be dead, diminished' had a lexicalised causative in PTK *\*ka-mate*, meaning 'to participate in the death watch, to sit with someone as they die'. The other languages with a causative do not allow such a meaning and focus on meanings centring around 'kill, murder' and so forth. The PTK meaning probably arose as an extreme extension of the 'facilitate' meaning which appears to have been current with the causative in PMC.

PMC *\*matakū* 'to be afraid' combined with the extended use of the 'up' directional, *\*sake*, does not always impart the expected inceptive meaning ('to get scared') seen in MRS, KSR and PON. Rather, there is also a durative meaning of 'to be a coward'.

## 5. DISCUSSION

### 5.1 THE DISTRIBUTIONS IN RELATION TO JACKSON (1983)

Jackson spoke of PMC, as I have in this paper, as if Nauruan did not exist. There is simply too little known about it. So when Jackson's conclusions or my own are mentioned in reference to PMC, we are really speaking of PMC without consideration of Nauruan. Jackson (1983:449-460) dissected the data available on Micronesian languages and presented a comprehensive subgrouping hypothesis for Nuclear Micronesian at a time when many of us working with him wondered if that would ever be possible. Some might say the evidence he developed is thin but I expect that will always be true, considering how rapidly the population appears to have dispersed to areas where the languages could begin to develop independently.

A more enduring criticism might be made in reference to work such as that of Pawley and Green (1984), whereas Jackson has used a radiation model in an attempt to characterise what may have been more of a network-breaking disintegration of the proto-language. His model for the higher level divergences suggests fairly abrupt breaks in linguistic communities when they may have remained in contact in a manner more similar to what we observe for Trukic today. However, the innovations he discovered meet the standard criteria for developing subgrouping arguments, and substantial sets of conflicting agreements did not emerge from his comparisons. The radiation model is appropriately employed under such circumstances.

Roughly, Jackson found evidence for suggesting that Kosraean first diverged, followed by Kiribati and then Marshallese. This left the long-recognised Ponapeic/Trukic group which disintegrated relatively late in time compared with the period in which the other splits occurred. Enigmatic evidence also tied Ponapeic to particular members of Trukic.

The evidence from the present study agrees completely in showing a well-marked Trukic group and a Ponapeic/Trukic unit. One comparison is in agreement with a MRS/PON/PTK group. The other evidence at first appears contrary to Jackson's higher order groups: the systematic agreements of KSR/PON/PTK in the extended use of the directional suffixes. And there were quite a number of idiomatic agreements between MRS/KSR/PON/PTK and the partial agreement of MRS with these others in showing vestiges of their extended directional suffix usages.

However, we need only postulate that KIR once had the extended directionals, but lost them, and say that KIR has become less active in allowing idiomatic deployment of the causative to dismiss those similarities as vestiges of PMC features, rather than a special, shared history of development between the languages other than KIR. I see nothing in the present data patternings to contradict Jackson's subgrouping hypothesis.

## 5.2 A SETTLEMENT AND LANGUAGE INTERACTION MODEL FOR THE DATA

Language distribution and migration theory seems well integrated into the work of most linguists studying the Oceanic Austronesian dispersal problem. It is probably still best formalised by Dyen (1956) and its basic premise is that a theory of language distribution is most likely to be correct when it has the fewest number of population movements occurring over the shortest distances. In a few final paragraphs that were stipulated to be very speculative, Jackson (1983) suggested a plausible settlement or dispersal model for the community of Nuclear Micronesian speakers. It involved Kosrae as the locality where PMC was spoken but is otherwise identical to what I propose below.

I do not find Kosrae a very satisfying homeland possibility, mainly for demographic reasons. I believe that the distances between Kiribati, the Marshalls, Kosrae and Ponape are small enough to have affected the settlement pattern of them. Thus, out of those which were settled first, that which had the greater population when another was discovered was most likely to have settled it and to have dominated the language that ultimately developed there. This reasoning gets rather stretched in reference to Kiribati and anything other than the Marshalls, but it is certainly true that distances between the Marshalls, Kosrae and Ponape are not so very great. And actually, the Marshalls are closer to Ponape than to Kosrae when going by way of Eniwetok and Ujelong.

What I would suggest is that an early population speaking PMC was spread through the Marshalls and Kiribati. Kosrae was then discovered and settled from the eastern atolls as a whole. Then early Marshallese would have begun to diverge from early Kiribati enough to leave the markings shared by MRS/PON/TRK that we observe today. The linear orientation of Kiribati running into the tighter spread of the Marshall atoll chains would have been a logical breaking point for political units and the beginnings of an ethnic boundary, which ultimately had linguistic consequences or vice versa.

Before early Marshallese was heavily marked as different from early Kiribati, Ponape would have been discovered, and settlement dominated by early Marshallese speakers would have ensued. By this model the Marshalls would have had an older, more developed population than Kosrae and been in a stronger position to dominate the language that developed on Ponape by having more people available for emigration. Early Kosraeans may also have had the lesser motive since there would be no qualitative improvement in environment, while early Marshallese speakers would have been moving from atolls to a high island with relatively abundant natural resources.

It is not difficult to imagine early Kiribati, Marshallese, Kosraean and Ponapean being mutually intelligible languages at this time. Conditions optimal for the spread of areal phenomena would have then existed. Pawley and Green (1984) would want to know if we appear to be dealing with a networking group of languages at this point. Grace (1981) would probably encourage us to broaden our database in a manner unaffected by standard descriptive biases and attempt to describe what we find in its own terms.

What we appear to be dealing with is an eastern area where some languages shared a few phonological and syntactic changes and may have calqued phrases or directly borrowed words and phrases from other languages or dialects. Word-final vowel devoicing may be an areal phenomenon that ran through this purported community of PON/KSR/MRS/KIR with KIR being the most remote and resisting it. The extended use of the directionals may be another such case. Some of the phrases I have shown in this work may have been calqued through the area. Jackson suggests numerous borrowings through the area, especially into Kosraean from the others. Bender's (1981) description of a mainly KIR/MRS distribution may have its roots in this period of time. But ultimately the languages did become quite distinct. Truk would be seen to have been settled from Ponape and the

influences of these languages on the others would have settled down to borrowing, except in the language network world of Trukic.

Jackson has shown that Kosraean is quite complex in relation to the other languages in terms of doublets, secondary patterns of POC reflexes, identifiable borrowings and other evidence of external influences. Even Kiribati, with its long recognised development under the influence of Polynesian, shows little such influence relative to the disturbances apparent in KSR.

I believe that some of the borrowings Jackson considered to be Trukic loans may have come from MRS, PON or KIR into Kosraean or could have been KSR loans into Trukic. His treatment of the labial stops and fricative (Jackson 1983:326-331) in particular seems to assume that changes in all the languages from PMC to their present phonetic configurations occurred at the moment of their respective divergences. Actually, they occurred over time, and interstage developments could have provided moments at which loans into KSR from MRS, KIR or PON could have the appearance of coming from modern Trukic or loans into Trukic from KSR might seem more easily explained on the basis of the current phonologies by positing the opposite.

It does seem, however, that some of the borrowings may have come from Trukic and along with an apparent Polynesian influence (Jackson 1983), Kosraean's history is very complex indeed. After fifteen to twenty years of general work on these languages and intensive work in one or more of them, most of us still wonder at the little we have learned about the synchronic phonological situation in KSR, let alone the diachronic processes that have brought it to its present state.

As I worked through Jackson's materials I kept wondering what social processes might have been responsible for Trukic loans into KSR or vice versa. The nearness of the KIR, MRS and PON languages allows a reasonable expectation of mutual influences between these and KSR but the borrowing may mainly be into KSR rather than out of KSR into the others (Jackson 1983) and the possible loans from TRK are very puzzling, given the distance.

Oral histories for Kosrae are almost non-existent but those of Ponape are rife with references to influence from Kosrae. Trukese oral history is similarly dominated by tales of immigrants and conquerors from *Achaw*. *Achaw* can simply mean 'basalt' in Trukic and there was an *Achaw-peiti* and an *Achaw-peitá* in these Trukese legends, Ponape and Kosrae, presumably. Kosrae was apparently the more prestigious as there is, through Trukic culture, a kind of process of implying that one knows more about Kosrae than others or has more Kosraean 'blood' than other people. The point is, so far as I know, there is no comparable claim by Ponapeans or Trukese to have run off to make war on Kosrae or to have had other contact that would have influenced the development of its language.

Perhaps any KIR or MRS influences on the development of KSR were mainly through emigration to Kosrae while the Ponapeic influences may have been obtained through regular social or economic contacts by way of Pingelap and Mokil. But the possibility of Trukic loans would suggest some form of extraordinary social or political relationship. I've wondered if the Kosraeans took slaves out of Truk. I've wondered if Kosrae was a kind of cultural centre where Truk and the others maintained emissaries to Kosraean royalty. I've wondered if Kosrae had emissary settlements on Truk (the Trukic loans then being introduced through bilingual children or something of the sort). What I have difficulty imagining is introduction of Trukic loans into Kosraean by conquest. Even that is possible, though, given a moment when Kosraean political organisation was fragmented and Trukese was centralised (the opposite of what was observed at the time of European contact).

## 6. EVALUATION OF THE PHRASE COMPARISON METHOD

This study produced more in the way of interesting materials to compare to external evidence than it might have in terms of internal comparison. I believe that I set out to do the right thing, in terms of developing a new line of evidence for these language relationships, but that I did so with the wrong kind of data. The choice of verbs in combination with the causative prefix and extended directionals meant that the apparent loss, in Kiribati, of extended uses of the directionals and a reduction in the idiomatic usages allowed with the causative obscured the significance of some otherwise interesting data. Similar problems dominated the comparison of other language data with Marshallese.

Were I to do it again I would develop a base of idiomatic phrases out of those occasionally given in the various dictionaries and make a comparison of the other languages against those materials. It would seem less likely to produce the problems encountered here due to employment of comparisons based on systematic grammatical processes. The comparison of phrases does, however, seem to be a productive line of inquiry. This study has shown that idiomatic phrases are quite durable in the grammars; they make these languages appear much more similar than they had seemed to me previously.

## APPENDIX: Data for the reconstructions in the order in which they are presented in the text

Notes: As in the text, MRS appears in both the common orthography which is given in quotations and in Bender's phonemic representations which are not. Voiceless vowels are in capital letters in WOL. In PTK the posited voiceless vowels have no special markings in this work.

*\*in(e,V)-ni-eti* 'oriented towards'

KIR	<i>inneti</i>	Vi, ( <i>in n eti</i> ), to lie or range in a straight line
PON	<i>inene</i>	Vt, to be oriented towards
	<i>inen</i>	adj, straight
	<i>inen-</i>	Vi, to come or go directly
WOL	<i>inney</i>	(as an exemplary Trukic language)
		<i>I inn ngali taatl.</i> I'm facing the sea.
		<i>I inney ngali taatl.</i> I've a clear view of the sea.

*\*ka-in(e,V)-ni-eti* 'to point something at something/someone'  
(or, perhaps, PMC *\*faka-in(e,V)-ni-eti*)

KIR	<i>ka-inneta</i>	Vt, to set in a straight line
MRS	<i>kijjiye-</i>	n., var., <i>kajjiye-</i> , in line with (dubious cognate)
PON	<i>ka-inene</i>	Vt, to straighten, inventory, aim
WOL	<i>ga-nney</i>	(as an example of Trukic)
		<i>I ga-nney ttowu-i iigA we</i>
		1SG CAUS-aim spear-my fish DEF
		I aimed my spear at the fish.



*\*rongo*

KIR	<i>ongo.</i>	to hear	
KSR	<i>El lohng.</i>	He (can) hear.	
MRS	<i>'I rong'.</i>	I (can) hear.	
PON	<i>I rong.</i>	I (can) hear.	
TRK	<i>U rong.</i>	I (can) hear.	
HAL	<i>U rong.</i>	I (can) hear.	
SAT	<i>I rong.</i>	I (can) hear.	
CAR	<i>I rong.</i>	I (can) hear.	
KSR	<i>Nga lohng kom.</i>		I heard you.
MRS	<i>'Ijrongjake yok'.</i>		I heard you.
PON	<i>I rongu-uk.</i>		I heard you.
PTK	<i>*I rongorong mwalia-mwu.</i>		I heard you (lit. I heard words-your).
TRK	<i>Ú rongorong mwánió-mw.</i>		I heard you (lit. I heard words-your).
HAL	<i>Ú rongorong mwálió-mw.</i>		I heard you (lit. I heard words-your).
SAT	<i>I rongorong mwálió-mw.</i>		I heard you (lit. I heard words-your).
CAR	<i>I rongorong mwálió-mw.</i>		I heard you (lit. I heard words-your).
WOL	<i>I rongorongO mwalia-mwU.</i>		I heard you (lit. I heard words-your).

*\*ka-rongorong (PTK)*

PTK	<i>*E pwe ka-rongorong-a Tony.</i>	He's going to make Tony overhear.
TRK	<i>E pwe arongorong Tony.</i>	He's going to make Tony overhear.
HAL	<i>E pwe arongorong Tony.</i>	He's going to make Tony overhear.
SAT	<i>E pwe arongorong Tony.</i>	He's going to make Tony overhear.
CAR	<i>E bwe arongorong Tony.</i>	He's going to make Tony overhear.
WOL	<i>E be garongorong Tony.</i>	He's going to make Tony overhear.

*\*ka-maturu (\*faka-maturu, perhaps, in PMC)*

KSR	<i>ahk-mutul</i>	(a) to put someone to sleep (b) to sneak up on fish, birds or women as they sleep
MRS	<i>'ka-majir'</i> <i>'an-ka-majir'</i>	to make someone sleep to sneak up on fish, birds or women as they sleep
PON	<i>E ka-mair seripein.</i> <i>I ka-mair mwaamw o.</i> <i>I paan ka-mair lii o.</i>	She put the girl to sleep. I'm going after the sleeping fish. I'm going to get that woman (sexually) as she sleeps.
TRK	<i>Ú pwe é-m-méúruw ewe semirit.</i> (Lagoon Trukese doesn't have the 'sneak' meaning)	I'm going to put that baby to sleep.
SAT	<i>I pwe a-m-maúruw kóókó we.</i> <i>I pwe a-m-maúr iik.</i>	I'm going to put the baby to sleep. I'm going to sneak up on the fish.
CAR	<i>I bwe a-m-maúruw ghóóghó we.</i> <i>I bwe a-m-maúr schóóbwut.</i>	I'm going to put the baby to sleep. I'm going to get a woman while she sleeps.
WOL	<i>I be ga-m-masúruwA góógóó we.</i> <i>I be ga-m-masúruw malúg.</i>	I'm going to put the baby to sleep. I'm going to steal the chickens while they sleep.



*\*mauru* (PTK *\*manawa*)

KSR	<i>El mouI.</i>	He's alive.
MRS	<i>E mewir.</i>	He's alive.
PON	<i>E mour.</i>	He's alive.
PTK	<i>*E manawa.</i>	He's alive.
TRK	<i>E manaw.</i>	He's alive.
HAL	<i>E menaw.</i>	He's alive.
SAT	<i>E menaw.</i>	He's alive.
CAR	<i>E melaw.</i>	He's alive.
WOL	<i>E melawA.</i>	He's alive.

*\*mauru-la*

KSR	<i>El pan mouI-lah.</i>	She had her baby (and all is well).
MRS	<i>'E mour-lok'.</i>	He's better (after illness).
MRS	<i>'E ka-mour'.</i>	She had her baby (and all is well).
PON	<i>E mour-la.</i>	She had her baby (and all is well).
PTK	<i>*E manawa-lako.</i>	She had her baby (and all is well).
TRK	<i>E manawa-nó.</i>	She had her baby (and all is well).
HAL	<i>E menawa-la.</i>	She had her baby (and all is well).
SAT	<i>E menawa-ló.</i>	She had her baby (and all is well).
CAR	<i>E melawa-ló.</i>	She had her baby (and all is well).
WOL	<i>Ye melawa-lagO.</i>	She had her baby (and all is well).

*\*mauru-sake*

MRS	<i>'E mour-tak'.</i>	He's getting better. It's erect (of a penis).
KSR	<i>El mouI-yak.</i>	He came back to life/revived. It's erect (of a penis).
PON	<i>E mour-da.</i>	see KSR glosses
PTK	<i>*E manawa-sake.</i>	see KSR glosses
TRK	<i>E manawe-tá.</i>	see KSR glosses
HAL	<i>E menawa-tá.</i>	see KSR glosses
SAT	<i>E menawa-tá.</i>	see KSR glosses
CAR	<i>E melawa-tá.</i>	see KSR glosses
WOL	<i>Ye melawa-takE.</i>	see KSR glosses

*\*ka-mauru* (PMC *\*faka-mauru*, perhaps)

KSR	<i>Nga ahk-mouI-yacl Joe me Camillus.</i>	I sustained Joe and Camillus.
MRS	<i>'I ar ka-mour Joe im Camillus'.</i>	I sustained Joe and Camillus.
PON	<i>I ka-mour Joe oo Camillus.</i>	I sustained Joe and Camillus.
PTK	<i>I ka-m-manawa-ara Joe me Camillus.</i>	I sustained Joe and Camillus.
TRK	<i>Ú a-m-manawa-ar Joe me Camillus.</i>	I sustained Joe and Camillus.
HAL	<i>Ú a-m-menawa-ar Joe me Camillus.</i>	I sustained Joe and Camillus.
SAT	<i>I a-m-menawa-ar Joe me Camillus.</i>	I sustained Joe and Camillus.

CAR	<i>I a-m-melawa-arJoe me Camillus.</i>	I sustained Joe and Camillus.
WOL	<i>I ga-m-melawa-arA Joe me Camillus.</i>	I sustained Joe and Camillus.

*\*ka-mauru-la* (PMC *\*faka-mauru-la*, perhaps)

KSR	<i>Nga ahk-moul-yacl-lah Joe ac Camillus.</i>	I saved Joe and Camillus.
MRS	<i>'I ar ka-mour-lok Joe im Camillus'.</i>	I saved Joe and Camillus.
PON	<i>I ka-moura-la Joe oo Camilus.</i>	I saved Joe and Camillus.
PTK	<i>*Ika-manawa-ara-lako Joe me Camillus</i>	I saved Joe and Camillus.
TRK	<i>Ú a-manawa-ara-nóJoe me Camillus.</i>	I saved Joe and Camillus.
HAL	<i>Ú a-menawa-ara-le Joe me Camillus.</i>	I saved Joe and Camillus.
SAT	<i>I a-menawa-ara-lóJoe me Camillus.</i>	I saved Joe and Camillus.
CAR	<i>I a-melawa-ara-lóJoe me Camillus.</i>	I saved Joe and Camillus.
WOL	<i>I ga-melawa-ara-lakO Joe me Camillus.</i>	I saved Joe and Camillus.

*\*mate*

KSR	<i>El mas.</i>	It's dead/diminished.
KIR	<i>E mate.</i>	It's dead/diminished.
MRS	<i>E mij.</i>	It's dead/diminished.
PON	<i>E mee-la.</i>	It's dead/diminished.
PTK	<i>*E mate.</i>	It's dead/diminished.
TRK	<i>E má.</i>	It's dead/diminished.
HAL	<i>E má.</i>	It's dead/diminished.
SAT	<i>E má.</i>	It's dead/diminished.
CAR	<i>E má.</i>	It's dead/diminished.
WOL	<i>E masE.</i>	It's dead/diminished.

*\*mate-la*

KSR	<i>Eltahl mas-lah ke smallpox.</i>	They're dying of/from small pox.
KIR	<i>mate-nako</i>	hopelessly adrift at sea
MRS	<i>'Rej mij-lak wot'.</i>	They're dying one after another.
PON	<i>E mee-la.</i>	He died/he's dead.
PTK	<i>E tamate-lako.</i>	He died/he's dead.
TRK	<i>A-a máá-nó.</i>	He died/he's dead.
HAL	<i>A-a máá-la.</i>	He died/he's dead.
SAT	<i>A-a máá-ló.</i>	He died/he's dead.
CAR	<i>A-a máá-ló.</i>	He died/he's dead.
WOL	<i>Ye sa mase-lagO.</i>	He died/he's dead.

*\*ka-mate* (PMC *\*faka-mate*, perhaps)

KSR	<i>John el ahk-misac-ilyac Tom.</i>	John murdered Tom.
KIR	<i>ka-mate-a</i>	kill, slay
MRS	<i>Ej ka-mej John.</i>	He paralysed John.
MRS	<i>Ej ka-m-mej John.</i>	He/it kept John awake.
PON	<i>I ke-mee-la John.</i>	I killed John.

PTK	* <i>Ita ka-mate-a tama-i.</i>	I stayed with my father as he died.
TRK	<i>Ú a a-máá-y semá-y.</i>	I stayed with my father as he died.
HAL	<i>Ú a a-máá-y semá-y.</i>	I stayed with my father as he died.
SAT	<i>I a a-máá-y semá-y.</i>	I stayed with my father as he died.
CAR	<i>I a a-máá-y semá-y.</i>	I stayed with my father as he died.
WOL	<i>I sa ge-masey tema-i.</i>	I stayed with my father as he died.

## BIBLIOGRAPHY

- BELLWOOD, Peter, 1978, *Man's conquest of the Pacific: the prehistory of Southeast Asia and Oceania*. Auckland: William Collins.
- 1984, The great Pacific Migration. *Encyclopaedia Britannica: Yearbook of science and the future*, 80-93. Chicago: Encyclopaedia Britannica.
- BENDER, Byron, 1981, A fossilized article in Marshallese. In Hollyman and Pawley, eds 1981:209-228.
- BLUST, Robert, 1984, Malaita-Micronesian: an Eastern Oceanic subgroup? *Journal of the Polynesian Society* 93:99-140.
- DYEN, Isidore, 1956, Language distribution and migration theory. *Language* 32:611-626.
- 1965, A lexicostatistical classification of the Austronesian languages. *International Journal of American Linguistics*, Memoir 19.
- GRACE, George W., 1955, Subgrouping of Malayo-Polynesian: a report of tentative findings. *American Anthropologist* 57:337-339.
- 1981 *An essay on language*. Columbia, South Carolina: Hornbeam Press.
- GREEN, Roger C., 1976, Lapita sites in the Santa Cruz Group. In Green and Cresswell, eds 1976:245-265.
- 1981, Location of the Polynesian homeland: a continuing problem. In Hollyman and Pawley, eds 1981:133-158.
- GREEN, Roger C. and M.M. CRESSWELL, eds, 1976, *Southeast Solomon Islands cultural history: a preliminary survey*. Royal Society of New Zealand Bulletin 11. Wellington: Royal Society of New Zealand.
- GREEN, Roger C. and J. MITCHELL, 1983, New Caledonia culture history: a review of the archaeological sequence. *New Zealand Journal of Archaeology* 5:19-67.
- HOLLYMAN, Jim and Andrew PAWLEY, eds, 1981, *Studies in Pacific languages & cultures, in honour of Bruce Biggs*. Auckland: Linguistic Society of New Zealand.
- IRWIN, G.J., 1980, The prehistory of Oceania: colonisation and culture change. In *The Cambridge Encyclopaedia of Archaeology*, 324-332. Cambridge University Press.
- JACKSON, Frederick H., 1983, The internal and external relationships of the Trukic languages of Micronesia. PhD dissertation, University of Hawaii at Manoa, Honolulu, Hawaii.
- KENNEDY, J., 1982, Archaeology in the Admiralty Islands: some excursions. *Bulletin of the Indo-Pacific Prehistory Association* 3:22-35.
- 1983, On the prehistory of Western Melanesia: the significance of new data from the Admiralties. *Australian Archaeology* 16:115-122.
- LEE, Kee-dong, 1974, Verbal aspect in Kusaiean and Ponapean. Working Papers in Linguistics. University of Hawaii 5/9:23-66.
- NATHAN, Geoffrey S., 1973, Nauruan in the Austronesian Language Family. *Oceanic Linguistics* 12:479-501.
- n.d., A grammatical sketch of Nauruan. MS.
- PAWLEY, Andrew, 1981, Melanesian diversity and Polynesian homogeneity: a unified explanation for language. In Hollyman and Pawley, eds 1981:269-309.
- 1982, Rubbish-man commoner, big man chief? Linguistic evidence for hereditary chieftainship in Proto-Oceanic Society. In Siikala, ed. 1982, 33-52.
- PAWLEY, Andrew and Roger C. GREEN, 1973, Dating the dispersal of the Oceanic languages. *Oceanic Linguistics* 12:1-67.

1984, The Proto-Oceanic language community. *The Journal of Pacific History* 19/3:123-146. Also in R. Kirk and E. Szathmary, eds *Out of Asia: peopling the Americas and the Pacific*, 161-184. Canberra: Department of Pacific History, Australian National University.

**SHUTLER, Richard, Jr and Jeffrey C. MARCK**, 1975, On the dispersal of the Austronesian horticulturalists. *Archaeology and Physical Anthropology in Oceania* 10:81-113.

**SIKALA, Jukka**, ed., 1982, *Oceanic studies: essays in honour of Aarne A. Koskinen*. Helsinki: Suomen Antropologinen Seura (The Finnish Anthropological Society).

THE UNIVERSITY OF CHICAGO  
DIVISION OF THE PHYSICAL SCIENCES  
DEPARTMENT OF CHEMISTRY  
5708 SOUTH CAMPUS DRIVE  
CHICAGO, ILLINOIS 60637  
TEL: 773-936-3700  
FAX: 773-936-3700

## TANIMBAR-KEI: AN EASTERN INDONESIAN SUBGROUP

ROGER F. MILLS

### 1. INTRODUCTION<sup>1</sup>

Three languages of eastern Indonesia appear to form a distinct subgroup: Fordata and Yamdena, spoken in the Tanimbar Islands, and Kei, in the Kei Islands. We owe our knowledge of these languages to the Roman Catholic missionaries Drabbe and Geurtjens, who worked in the early decades of this century. They seem to have had an implicit understanding of the relationship between the languages, but they were not comparativists, and their work was not oriented in that specific direction. Fortunately, the dictionaries and grammars they left us are of very high quality (Geurtjens's peculiar spelling system notwithstanding), and provide excellent foundations for comparative work.

The first modern discussion of these languages is found in Dyen (1965), where on the basis of 'adequate' lists Kei is ungrouped within the Moluccan linkage, its highest percentage (20.9) with Leti. Unfortunately, the two Tanimbar languages were represented by 'subadequate' lists, but showed higher percentages: Fordata-Kei 37.6%, Fordata-Yamdena 31.5%. Dyen (p.45) goes on:

*Fordat* forms the Fordatic subgroup with Kei...Whether Fordatic is a closed group depends on Fordat's percentage with Jamden. If this percentage is reliable... Jamden is a member of the same subgroup.

Yamdena, he says, "may be no more closely related to Fordat than to Tettum (27.3%)". (Presumably Yamdena-Kei scored even lower than that.) But the resemblances between Fordata and Yamdena, and their divergences with Tettum, are, as we like to say, obvious upon inspection. And frankly, the percentages are not reliable, at least in part because of very conservative judgments as to cognacy,<sup>2</sup> even though in a relative way a correct picture of the relationship happens to emerge.

---

<sup>1</sup>To my regret, I have not been privileged to know Professor Grace except through his many scholarly contributions. It is nonetheless a great honour to present my work to him on this occasion. Considerations of space have forced upon me an unaccustomed brevity. Thus, some of the presentation may seem quite summary – but it is certainly not without ample supporting evidence, almost all of which is readily available in well-known publications. If some of my claims seem to require leaps of faith, be assured I do not mean to be dogmatic. The much longer draft from which this paper is condensed is available to anyone ineresed. I wish to express my appreciation to the University of Michigan's Center for South and Southeast Asian Studies for its assistance and logistical support over several years, without which my research would have been much impeded.

<sup>2</sup>Given the availability of good dictionaries, whose several thousand entries afford ample data for detailed comparison, I have never seen the virtue of sorting out just 200 words from each. Happily, two recent Indonesian studies (Tetelepta et al. 1985; Pieter et al. 1986) contain word lists recorded from informants, from which a 'subadequate' Yamdena-Kei list could be compiled. It contained just 125 out of 197 items but, of these, 61 (or 49.2%) were clearly cognate. Given that Yamdena and Kei are the most distantly related of the three, Fordata-Kei must surely score higher than Dyen's 37.6%.



It is also probably true that Tanimbar-Kei and Tettum ultimately subgroup together, but at what level is not clear. I hope in this paper to offer a tentative answer to that question, but limitations of space preclude detailed treatment.

A more recent discussion, based on fieldwork and a "cursory examination of Geurtjens and Drabbe", is found in Collins (1982). He considers Kei and Fordata to be members of a "far-flung dialect chain". Since the published sources describe only the two extremes of this putative chain, it is quite possible that intervening areas provide the missing links, but in my opinion Fordata and Kei are now clearly distinct languages (see section 4.2).

## 2. NEIGHBOURS

Drabbe (1932a) reports on a third Tanimbar language, Selaru. Its vocabulary seems to show strong resemblances to Yamdena, but the word list is short and one has the impression that Drabbe's choices may have been influenced by his greater command of Yamdena. As will be shown, on phonological grounds Selaru is less closely related to Fordata, Kei or Yamdena than they are to each other.

Yet a fourth Tanimbar language, Selwasa, is brought to light in Collins (1982); it is mentioned by Coward (1989:2) as being closely related to Selaru ("at 56% lexical similarity").

A glance at a map shows that Tanimbar and Kei are both quite large and compact island groups; Kei lies some  $\pm 200$ km north-north-east of Tanimbar. To the east ( $\pm 200$ km) their only neighbour is the Aru group, whose languages are evidently Austronesian but very divergent (see van Eijbergen 1864; Collins 1982 tentatively groups Aru with Tanimbar-Kei). West of Tanimbar ( $\pm 150$ km) the nearest large island is Babar and environs, whose known languages belong to the Lettic group.<sup>3</sup> Between Kei and the eastern tip of Seram lie several small islands. Tetelepta et al. (1985) mention Kur and Teor as languages spoken in that area but not subgrouping with Kei; the brief discussion in Collins (1982:122-123) confirms this view. On Kei itself, a language called Eli-Elat (after the two villages where it is spoken) is said to be the survivor of the original language of the Banda Islands, whose inhabitants were dispersed in the seventeenth century. (See Stresemann 1927:182-190; he held that Eli-Elat did not subgroup with his Ur-Ambon, nor does it appear to subgroup directly with Tanimbar-Kei. Collins, again, confirms this view.)

## 3. BRIEF PHONOLOGICAL HISTORY OF TANIMBAR-KEI

Comparison of Fordata, Kei and Yamdena allows the fairly straightforward reconstruction of Proto Tanimbar-Kei (PTK) with the following sound system:<sup>4</sup>

Stops: *-p(?)*, *b*, *mb*, *t*, *d*, *nd*, *k*, *ŋg*

Nasals: *m*, *n*, *ŋ*

Resonants: *w*, *r*, *l*, *y*

<sup>3</sup>That is, Wetan, so well reported in Josselin de Jong (1987, researched in the early 1930s), and 37 words of a language called "Tanimber (Baba)" recorded in Earl (1848), which shows resemblances to both Wetan and Kisar.

<sup>4</sup>Abbreviations used are as follows: AMB – Ambonese, AN – Austronesian, CMP – Central Malayo-Polynesian, EIN – Eastern Indonesian, EMP – Eastern Malayo-Polynesian, F – Fordata, GAL – Galoli, K – Kei, KIS – Kisar, LET – Leti, LS – Luang-Sermata, MN – Melanesian, OC – Oceanic, PBD – Proto Barat Daya, PBDS – Proto Barat Daya-Selaru, PCMP – Proto Central Malayo-Polynesian, PLet – Proto Lettic, PLS – Proto Lettic-Selaru, PSEM – Proto SE (Southern) Maluku, PTK – Proto Tanimbar-Kei, ROT – Roti, SHWNG – South Halmahera West New Guinea, SLAR – Selaru, TET – Tettum, TIM – Timorese, WET – Wetan, Y – Yamdena.

Voiceless continuants: *f, s*

Vowels: *i, e, a, u, o*

The PTK sound system reflects that of AN and CMP (at least as I conceive it) as set forth in Table 1. Crucial to the definition of PTK as a distinct entity are the merger of AN \**d, D* and \**z, Z* > PTK \**d*, the merger of AN \**r, j, R* > PTK \**r*, the loss of AN \**-w* and \**-y*, and the treatment of the vowel sequences that arose after loss of the AN/CMP laryngeals. Some of the reflexes in the Table require further comment:

(a) PTK \**-p* and \**b* are needed to account for Yamdena *-p* and *b*.<sup>5</sup> Their presence here implies their presence in CMP, contrary to Stresemann's position but in line with more recent views (e.g. Collins 1980; Blust 1981). Though rare, stop reflexes of AN \**p* and \**b* are found in eastern Indonesian languages; their occurrence in Yamdena is indicative of its overall conservatism.

TABLE 1: AN-CMP-PTK REFLEXES

AN	<i>p</i>	<i>b</i>	<i>mp, mb</i>	<i>m</i>	<i>t, T, C</i>	<i>s, c</i>	<i>n</i>	<i>ñ</i>	
CMP	<i>p</i>	<i>b</i>	<i>mb</i>	<i>m</i>	<i>t</i>	<i>s</i>	<i>n</i>	<i>ñ (or n?)</i>	
PTK	<i>f-f-p</i>	<i>b</i>	<i>mb</i>	<i>m</i>	<i>t</i>	<i>s</i>	<i>n</i>		
AN	<i>d, D</i>	<i>z, Z</i>	<i>nt, nT, nd, nD, nz, nZ, nc, ns, nj</i>			<i>k</i>	<i>g</i>	<i>ŋk, ŋg</i>	<i>ŋ</i>
CMP	<i>d</i>	<i>Z</i>	<i>nd (+ possible others)</i>			<i>k</i>	<i>g</i>	<i>ŋg</i>	<i>ŋ</i>
PTK	<i>d</i>		<i>nd</i>			<i>k</i>	<i>ŋg</i>		<i>ŋ</i>
AN	<i>r</i>	<i>j</i>	<i>R</i>	<i>l</i>	<i>w</i>	<i>y</i>	<i>q, S, ?, ∅</i> etc.		
CMP	<i>r</i>	<i>j</i>	<i>R</i>	<i>l</i>	<i>w</i>	<i>y</i>	<i>q, S, ∅</i>		
PTK	<i>r</i>			<i>l</i>	<i>w-w-∅</i>	<i>y-y-∅</i>	<i>∅</i>		
AN	<i>i</i>	<i>u</i>	<i>ə<sub>1</sub></i>	<i>a</i>	<i>ə<sub>2</sub></i>	<i>-ay, -əy</i>	<i>-aw</i>	<i>-uy</i>	<i>-iw</i>
CMP	<i>i</i>	<i>u</i>	<i>ə, e, o</i>	<i>a</i>	<i>ə, e, o</i>	<i>-ay</i>	<i>-aw</i>	<i>-uy</i>	<i>-i?</i>
PTK	<i>i</i>	<i>u</i>	<i>a</i>		<i>e</i>	<i>-a</i>		<i>-u</i>	<i>-i</i>
AN	<i>a(qS∅)i</i>		<i>a(qS∅)u</i>						
CMP	<i>a(qS∅)i</i>		<i>a(qS∅)u</i>						
PTK	<i>e</i>		<i>o</i>						

(b) PTK \**nd* (like Stresemann's Ambonese \**nd*) in fact clearly reflects only AN \**nt*, \**nd/D* and perhaps \**ns*; examples for the other cases are either doubtful or altogether lacking. Collins (1980) gives evidence that at least \**ns* shows distinct reflexes in some Ambonese languages.

(c) The merger of AN \**r, j, R* > \*\**r* is also seen in Lettic; Selaru has *r* < \**r, R*, but contradictory evidence for \**j*.<sup>6</sup>

(d) PTK \**y* is traceable to AN \**y* only in medial position. PTK \**y*- has multiple origins, among them (a) reduction of two homophonous prefixes \*\**i*- '1. personal article, 2. locative marker' and (b) 'y-accretion' preceding initial \*(*qS∅*)*a* (more frequent in Fordata and Kei than in Yamdena). This last is found in several EIN languages as well as in Fijian and perhaps other MN languages. But

<sup>5</sup>Yamdena also has a [p] in other positions. In certain cases, it is a conditioned variant of *mb*; in others, a free variant thereof (see Drabbe 1926b:4). This latter apparently reflects influence from the 'northern' dialect of Yamdena, where *p* < \**mb*.

<sup>6</sup>Collins (1982:127) claims SLAR *s* < \**j*; no example is cited, but I suspect he had in mind SLAR *waras* 'liana' < \**waRcj*. At least in Drabbe, there are no unambiguous witnesses for \**j*: we see *-∅* twice, *-∅-* twice, *-l-* once, and all can be explained. SLAR *waras* is probably < Yamdena *waras*, itself a likely loan (like *fase* 'rice') from Seran Laut (Geser) where \**j* > *s* is regular.

most instances of PTK *\*y-* are of uncertain origin. In no case does *\*y-* occur preceding *\*i* or *\*u*. These statements apply equally well to Lettic *\*y*.

(e) PTK *\*o* (except when < AN *\*a(qSØ)u*) is of obscure origin, as it is, I feel, in most EIN languages, aside from those few, like Tettum, where it reflects AN *\*ə*. It is likely due to fairly heavy influence from some outside source – languages like Gorontalo or Mongondow of North Sulawesi, or Banggai, Mori or Toraja of East/Central Sulawesi are likely candidates, but possible very early influence from a Proto Oceanic-type language would be a tempting hypothesis.

(f) Where AN *\*a(i)* and *\*a(u)* > PTK *\*e,o*, the resulting monosyllable is restored to bisyllabicity in one of two ways: in Fordata and Kei, by ‘lengthening’ with a following *-a-*; in Yamdena by addition of *-e (=a)* to the base (e.g. AN *\*baSu* > PTK *\*bo-n* > F *voa/n*<sup>7</sup> ‘smell’, K *woan*, Y *bone* ‘to kiss’). The treatment is the same for those rare monosyllables containing AN *\*ə* – *\*ənəm* ‘six’ > PTK *\*nem* > F *nean*, K *neyen*, Y *neme*. Thus nothing in these languages compels the reconstruction of a separate reflex of the AN sequences; that is not the case, however, for Lettic and other presumably related neighbours.

(g) Final consonants. PTK shows final *\*p t k m n ŋ r l s*. Yamdena generally retains these. The situation in Fordata and Kei is less clear, since all finals are liable to loss, depending on various factors: (a) inalienable nouns lose finals, replacing them with possessive markers; (b) many verbs lose their final, either outright or by replacement with *-t, -k, -n* or *-ŋ* (Yamdena also has occasional replacement of verbal finals). These ‘suffixes’ apparently had some as yet undetermined morphological value: at present only *-t* and *-k* can be traced back to higher levels; (c) *\*-p* and *\*-m*, if not otherwise lost, shift > *-t* and *-n*, and likewise *\*-ŋ* > *-n*; (d) *\*-s* >  $\emptyset$ , probably via F-K *\*\*h*, with rare *-s* likely borrowed; (e) also *\*-k*, via F-K *\*\*?*, >  $\emptyset$ . Thus in Fordata and Kei, only *-r, -l* unambiguously reflect original finals.

#### 4. DEVELOPMENTS FROM PTK

##### 4.1 YAMDENA

The phonological conservatism of this language has already been noted; it retains the PTK sound system with just two changes, that is, merger of *\*ŋg* and *\*ŋ* > *ŋ*, and surface realisation of *\*-a#* as *e* (with *-a-* before suffixes). It also retains traces of what were apparently allophones of PTK *\*e* (see section 4.1.1a). Other developments that distinguish Yamdena from Fordata and Kei are: infrequency of *y*-accretion; sporadic shift of ultima *\*u* > *i* (see section 4.1.1b); use of *\*\*a* to lengthen monosyllables; and a sizeable body of words with ‘*o* of obscure origin’ without F-K counterparts.

##### 4.1.1 TWO PECULIARITIES OF YAMDENA PHONOLOGY

(a) The allophones of PTK *\*e* were evidently a higher [e]/\_\_C{i,u} versus a lower [ɛ]/\_\_C{a} (i.e. elsewhere). Because of its frequency, [ɛ] was probably the norm. Drabbe’s ‘dictionary’ dialect maintains the distinction with written ‘*é*’ versus ‘*e*’, as does Geurtjens’s Kei, writing ‘*i*’ corresponding to Yamdena ‘*é*’ versus ‘*e*’ otherwise. Compare K *riŋ*, Y *réŋi* ‘carved figures’; K *tíl*, Y *téli* ‘three’; K *ref*, Y *defe* ‘fathom’. (There are discrepancies, some surely due to omission of the

<sup>7</sup>In this paper *a /* is used in examples to separate an irrelevant ending or portion of the word cited.

diacritics.) In Drabbe's Yamdena, the distinction has become phonemic, probably through sporadic lowering of *\*i > é/e* together with borrowing from languages whose /e/ norm was closer to [e] or [ɛ] in 'wrong' environments. Neither of the recent Indonesian studies, however, notes the distinction in Yamdena or Kei. It is certainly possible for it to have been lost in the past sixty-plus years, but in view of the overall quality of these works, one cannot rule out simple oversight.

(b) Shift of ultima *\*u > i* is observed in some 30 forms, several of them, like *téli* 'three' < *\*təlu*, 'basic' vocabulary. There is no apparent phonetic conditioning – note *féni* 'full' < PTK *\*fenu* versus *fenu* 'wear something around the neck' < AMB *\*ənu* 'neck', *manik* 'bird' < *\*manuk*. Several languages in the area show the same shift: Leti/Moa sporadically, regularly in Buli, Bonfia, various Ambonese languages and (Letic) Wetan of Babar which, being geographically closest to Tanimbar, is a likely source.

In contrast, Fordata shows *\*-u > i* in just a handful of forms, only a few of which match Yamdena. The implication is that the two languages received their *i*-forms from different sources or at different times.

#### 4.2 FORDATA AND KEI

These two languages clearly underwent a period of common development during which the following changes took place:

- (a) loss/replacement of final C, as discussed above;
- (b) *\*s > F-K \*\*h*;
- (c) *\*k > F-K \*\*?* (with later *\*\*h, ? > 0/\_#*);
- (d) (ordered) 1. *\*b,d > F-K \*\*v,r*; 2. *\*mb,nd > F-K \*\*b,d*;
- (e) (ordered) 1. *\*ŋ > F-K \*\*n*; 2. *\*ŋg > F-K \*\*ŋ*;
- (f) retention of PTK *\*e* allophones;
- (g) 'lengthening' of monosyllables in *\*e,o* by addition of *-a*;
- (h) Diphthongisation of final high vowels in the environment *\*\*{e,a} {h,?}\_(C)#*,

such that, for example, *\*\*-ahu > -ahaw* as in K *yaháw*, F *yáha* 'dog' < *\*asu*. There are questionable examples that suggest the environment may have included other vowels. This rule evidently began to operate about the time the two languages were separating, as it is seen sporadically and often irregularly in Fordata.

##### 4.2.1 KEI

From the stage just outlined, Kei exhibits further changes:

(a) loss of *\*\*?.?*. Although Geurtjens's spellings are unclear, and Tetelepta offers contradictory forms, *loss* seems the better assumption.

(b) loss of *\*1/{a,u}\_#* and intervocalically between any *a-u* sequence. I posit that *l* developed a velarised *w*-like allophone in these environments, which subsequently dropped/blended with

neighbouring *u*, for example, *n-a* < \**ala* 'give, take'; *u* ~ *uu-n* 'head' < \**ulu(-n)*; *wowt* = *wawt* 'image' < \**walut*; *wuan* 'moon' < \**bulan*.

(c) loss of final vowels. This affected original final *V*, as well as those exposed by loss of final *C* during the F-K stage. Final *V* loss evidently took place while 'l-loss' was on-going, but before \*\*[ɪ] was lost – vowels exposed by 'l-loss' are retained.

(d) reduction of *VV* sequences > diphthongs in the case of *V+* {*i,u*}, long *V* in the case of like vowels. This is one of the knottier areas of Kei phonology, since Geurtjens's orthography not only over-differentiates, but is also rather inconsistent. Tetelepta's 'phonemicization' seems more consistent, but is, I suspect, under-differentiated.

(e) Development of final stress – in fact a corollary of the predominantly monosyllabic structure of most bases.

#### 4.2.2 FORDATA

From the F-K stage, *F* shows only slight further change:

(a) loss of \**e* allophones;

(b) retention (development?) of penultimate stress;

(c) The 'dictionary' dialect shows sporadic loss of medial \*\**h*?, general loss of \*\**h*- (versus retention in the Molu dialect). Where \**s* > *h/e*\_\_*V*, the outcome in every case is *-ahi-*, as in *ahil* 'regret' < PTK \**sesal*; *fahi* 'fart' < PTK \**fesu*; *vahi* 'paddle', K *wehe*, AN \**bə(R)say*. In all cases I believe diphthongisation was responsible: \*\**ehV* > *ehe*Ÿ > *éhe*∅ > *éhi* (constraint on final *e*) > *-áhi-* (dissimilation).

We might hypothesise that Fordata failed to continue developing toward a Kei-like stage by contact with conservative Yamdena after moving into the Tanimbar area. Interestingly, however, there is clearer evidence for Fordata influence on Yamdena than vice-versa, for example, occasional *Y* *ff* for \**b*, likely < *F* *v*; *Y* ∅-∅- for \**s* surely < *F* *h*; *Y* *r* for \**d*, < *F* *r*.

### 5. SHARED GRAMMATICAL FEATURES

#### 5.1 PRONOUNS

With appropriate sound changes, the three languages have almost identical sets of personal pronouns, as seen in Table 2.

TABLE 2: FORDATA-KEI-YAMDENA PERSONAL PRONOUNS

	1SG	2SG	3SG	1PL.INC	1PL.EXC	2PL	3PL
PTK	* <i>yaku</i>	* <i>ko</i>	* <i>ia</i>	* <i>kita</i>	* <i>kami</i>	* <i>kimi</i>	* <i>sira</i>
F	<i>yaʔa</i>	<i>oa</i>	<i>ia</i>	<i>ita</i>	<i>ami</i>	<i>mia</i> †	( <i>h</i> ) <i>ira</i>
K	<i>yāw</i> , <i>yaʔaw</i>	<i>o</i>	<i>i</i>	<i>it</i>	<i>am</i>	<i>im</i>	<i>hir</i>
Y	<i>yaku</i>	<i>kou</i> (< ? <i>kowe</i> = / <i>ko+a</i> /)	<i>iye</i> (= / <i>ia</i> /)	<i>kite</i>	<i>kami</i>	<i>kmi</i> †	<i>sire</i>

†(Apocope in *F*, syncope in *Y*, unexplained. Perhaps *F* *mi+a* < \*\**mi* by influence of *Y* *kmi*.)

Table 3 displays the possessive forms. Yamdena shows some analogical extension of the *\*ni*-formative; Fordata and Kei have unexplained restructurings in their plurals.

TABLE 3: FORDATA-KEI-YAMDENA POSSESSIVES

(a) *Preposed, alienable forms.*

	1SG	2SG	3SG	1PL.INC	1PL.EXC	2PL	3PL
PTK	<i>*niggu</i>	<i>?*mu</i>	<i>?*ni</i>	<i>*(ni)nda</i>	<i>*mami</i>	<i>?</i>	<i>?</i>
F	<i>nig(u)</i>	<i>mu</i>	<i>ni</i>	<i>did(a)</i>	<i>mam(i)</i>	<i>bir</i>	<i>rir</i>
K	<i>nig</i>	<i>mu</i>	<i>ni</i>	<i>did</i>	<i>mam</i>	<i>bir</i>	<i>rir</i>
Y	<i>nigu</i>	<i>nime</i>	<i>nia</i>	<i>ninde</i>	<i>mami</i>	<i>mir</i>	<i>nir</i>

(b) *Suffixed, inalienable forms.*

F	<i>ŋ</i>	<i>m</i>	<i>n</i>	<i>d</i>	<i>m,mami</i>	<i>b</i>	<i>r</i>
K	<i>ŋ</i>	<i>m</i>	<i>n</i>	<i>d</i>	<i>b</i>	<i>b</i>	<i>r</i>
Y	<i>ŋ</i>	<i>m</i>	<i>n</i>	<i>nind(ar)</i>	<i>mamyar</i>	<i>mir</i>	<i>nir</i>

## 5.2 VERBAL PERSON-MARKING PREFIXES

In Fordata and Yamdena, the verbal personal prefixes have two forms: *long*, as cited in Table 4, and *short*, in which final *i,u* undergo 'binding' (see section 8), while *-a* drops out. The distribution of short versus long forms is not entirely clear; Drabbe cites cases of syntactic conditioning (long forms are emphatic), lexical conditioning (some verbs require the long form), as well as phonological conditioning (long forms precede #CC-).

TABLE 4: FORDATA-KEI-YAMDENA VERBAL PERSONAL PREFIXES

	1SG	2SG	3SG	1PL.INC	1PL.EXC	2PL	3PL
F	<i>u</i>	<i>mu</i>	<i>na</i>	<i>ta</i>	<i>ma</i>	<i>mi</i>	<i>ra</i>
K	<i>u</i>	<i>um</i>	<i>en~na</i>	<i>it</i>	<i>am(u)</i>	<i>im</i>	<i>er~hir</i>
Y	<i>ku</i>	<i>mu</i>	<i>na</i>	<i>ta</i>	<i>ma</i>	<i>mi</i>	<i>ra</i>

## 5.3 PLURALISATION

Fordata and Yamdena have a productive plural suffix, F *-r*, Y *-r ~ -ar*. It has been lost in Kei, but is seen fossilised in the possessives *bir* and *rir*.

## 5.4 DERIVATIONAL AFFIXES

Drabbe's analyses of Fordata and Yamdena derivational affixes (all prefixes) show that with minor exceptions the two systems are identical. Adjectival (stative, intransitive) prefixes include *ma-*, *ŋa-*, *ka-* in both and F *ba-*, Y *mba-*. Both show the causative prefix *fa-*. Further, F, Y *ra-*, *sa-*, *si-*, though described in rather vague terms, seem to have similar functions. The vowel of these prefixes usually



deletes when the (long) person markers are added, so (for example in third person singular) *nam-*, *naŋ-*, *naf-*, *nat-* etc.

For Kei, Geurtjens gives an equivalent list: stative *nam-*, *naŋ-*, *nak-*, *nab-*; causative *naf-*, *nat-* ~ *nad-* and *nas-*. There are variants written with 'ĕ' and 'e', possibly 'euphonic' or else lexically determined. Geurtjens's discussion of all these is disappointing and rather confused; it seems unlikely that morphemes of such frequent occurrence can truly be as 'meaningless' as he often says. On the other hand, Tetelepta's analysis is no more illuminating; perhaps Kei has indeed lost some of the grammatical or lexical motivation for its morphology.

### 5.5 BINDING

Fordata and Yamdena show the morphophonemic process I term 'binding'; Kei has only a remnant of it (see section 8).

## 6. A UNIQUE (?) PHONOLOGICAL INNOVATION

Based on the following:

- F roal* '(go) seaward (away from the speaker)'  
*K roa* 'the sea as a thoroughfare'  
*Y dole* 'seaward; easterly; foreign, not Yamdena'

we can construct PTK *\*dol* 'seaward', presumably by consonantal metathesis < *\*lod* < AN *\*laSud*. Fordata and Kei further show reflexes of *\*\*ndol* – *F doal* 'seawards (toward speaker)', *K en/do* 'come from the south or west; come or do something in the direction of speaker'. It cannot be determined whether *\*\*ndol* is a F-K innovation (by analogy with other *r ~ nd* forms?) or a reflex of PTK *\*\*ndol*, lost in Yamdena.

Only one other case of a metathesised reflex of *\*laSud* is known to me: Aru (Ujir dialect) *adohl* 'east' (van Eijbergen 1864), in all likelihood a borrowing from the Tanimbar-Kei area.<sup>8</sup>

At least one other language has a *\*laSud* reflex with similar directional senses: Buli *lalau* 'seaward, away from speaker', *malau* 'seaward, toward speaker'. It is possible that *la-* and *ma-* in these words reflect forms meaning 'go' and 'come'.

If this consonantal metathesis is indeed unique to PTK, it is an innovation that serves to define the subgroup. More interesting, perhaps, is what the form implies: in order for PTK to have *\*dol* < *\*\*lod* < *\*laSud*, the final *\*d* must have survived into the immediate ancestor of PTK, and so on back to CMP. A similar argument for CMP *\*-b* can be made on the basis of Leti *tutwu* 'to close, cover' reflecting CMP *?\*tutub*. The question whether voiced final stops were retained in CMP has not been researched, largely, I suspect, because most of us have assumed that devoicing of final stops would be a more expected development. The question is made more difficult, too, by the relative rarity of AN *\*-b* and *\*-d*, and the loss of finals in so many EIN languages. A handful of examples for *\*-b* in Lettic, and PTK *\*dol*, are the only evidence so far discovered.

<sup>8</sup>The metathesised form may have arisen in the first place by deliberate change in a special 'seafarers' language': one of those ritualised speech forms full of taboo-substitutes and circumlocutions often reported from various Indonesian areas (though not specifically from the T-K area). Such a 'seafarers' language' could well have taken this word to Aru.

## 7. THE POSITION OF SELARU

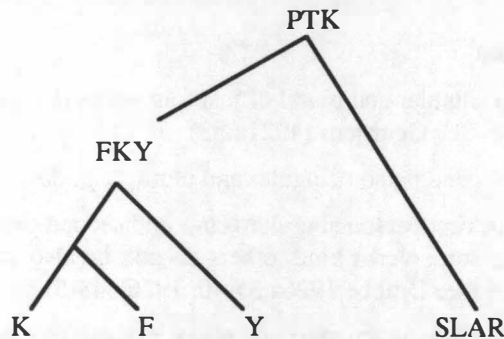
Selaru is spoken on the island of the same name at the southern end of the Tanimbar group. According to Coward (1989:2) there are also Selaru-speaking areas on Yamdena and some offshore islands. Coward estimates the number of speakers at 7000-9000, considerably more than Drabbe's estimate (1932a:1) of about 3600.

In view of its location, and the obvious cognacy of so many forms in Drabbe's word list, one might hastily conclude that Selaru is a close relative of Yamdena. I suspect, however, that Drabbe tended to concentrate on vocabulary that clearly resembled that of the (to him) more familiar Yamdena.

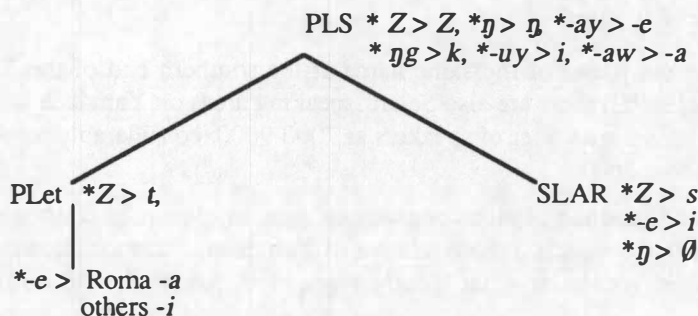
Like Yamdena, Selaru has the  $\acute{e}$ : $e$  distinction under the same conditions, but, as in Yamdena, with many 'wrong' occurrences; clearly some of the  $\acute{e}$ - or  $e$ -forms are borrowed from Yamdena. Coward notes the distinction, but inexplicably does not consider it phonemic. Clearly borrowed is Selaru *mb*, in a handful of forms almost all of which recur in Yamdena. Selaru has 'binding' in all possible environments, including two that Yamdena lacks – with the singular suffix *-ke*, and in third person singular of verbs, where the prefix is *i-* (unlike Y *n-*). Several phonetic changes, however, show that Selaru has had a quite different history than Fordata, Kei or Yamdena: AN  $*b >$  SLAR *h*, AN  $*p >$  SLAR  $\emptyset$  (with rare *f*, *h* probably borrowed), CMP  $*\eta g >$  SLAR *k*, AN  $*\eta >$  SLAR *n*,  $\emptyset/V\_V$ , SLAR *s <* AN  $*l/_\{i,u\}$ . Further, the base forms of CVCV nouns usually show loss of the final V, as do most verbs; with nouns, the final V reappears when suffixes are added (thus it is present underlyingly). Coward (1989) symbolises these finals with 'glide' phonemes – it remains unclear whether Drabbe's  $\emptyset$ -final, or Coward's glide-final forms can occur freely.

Two developments in particular show that any conclusion about 'close relationship' with Yamdena is indeed not only hasty but incorrect: Selaru has (1) *s <* AN  $*Z$  versus PTK  $*d$ , and (2) *-i <* AN  $*-ay/ey$ ,  $-uy$  versus PTK  $*-a$  and  $*-u$ . (Selaru like PTK has  $-a <$   $*-aw$ , but in just one example.)

One alternative is to amend PTK to include  $*Z$ ,  $*ay$  and  $*uy$  reflexes; Fordata, Kei and Yamdena still remain a discrete subgroup, but coordinate with Selaru, thus:



The other alternative is to align Selaru with some other group – but which? Its phonology compares well with that of Galoli (eastern Timor, Wetar I.), but yet another reflex is required for  $*R$  (GAL  $\emptyset$ , SLAR *r*). Further, since Galoli and Tettum are closely related, we should also need  $*\text{ə}$  (GAL, SLAR *e*, TET *o*). But this possibility is vitiated by lack of binding and productive pluralisation in Tettum and Galoli. The other local group where 'binding' and 'plural' are found is Lettic, which also has distinctive reflexes of  $*Z$  and  $*-ay/uy$ , with  $*-a <$   $*-aw$ . I therefore posit Proto Lettic-Selaru as follows:



As we will show later, Proto Letic-Selaru and Proto Tanimbar-Kei descend in turn from a common node – thus, all these south-eastern Indonesian languages which exhibit binding are related (as we might suspect in any case!).

## 8. BINDING

Binding is a morphophonemic process whereby at certain junctures (1) a final  $\{i,u\}$  of the first element and the initial C of the second element metathesise; (2) the shifted V reduces to  $\{y,w\}$ ; (3) the original V is deleted. (Before a V-initial base, the final  $\{i,u\}$  simply shifts to  $\{y,w\}$ .) Thus, the two elements are in a sense interlaced, bound together into a single entity. Very schematically:

$$\#...(C) \{i,u\} + (C)V... \implies \#...(C)\emptyset + (C)\{y,w\}V...$$

In an earlier paper (Mills and Grima 1980) this was termed ‘pseudometathesis’ in view of its phonetically explicable nature since it proceeds via two rules, (1) progressive assimilation of the vowel’s palatal or round quality, (2) V-deletion in pretonic, interconsonantal position.

Binding occurs actively to one degree or another in Fordata and Yamdena (only fossilised in Kei), Selaru and all Letic languages except Kisar. The relevant environments are as follows.

### 8.1 PERSONAL PREFIX + VERB

(a) Kei: in second person singular and plural of just four verbs (*ba* ‘go’, *mat* ‘die’, *ma* ‘come’, *an* ‘eat’; as in *um+ba* > *umbwa*). See Geurtjens (1921a:33).

(b) Fordata: regularly in second person singular and plural (e.g. *dava* ‘seek’, *mdwava*, *mdyava*).

(c) Yamdena: regularly in first person singular (*ku-*), and second person forms. In both Fordata and Yamdena, Drabbe notes, some verbs bind, others do not; he also cites verbs which apparently take both types of conjugation (see Drabbe 1926a:35-40; 1926b:48-51).

(d) Selaru: regularly in first person singular and plural exclusive (*mi-*), second person forms and third person singular (*i-*). Some verbs bind, others do not (Drabbe 1932a:7-10).

(e) Letic: Examples in Jonker (1932, *passim*) show binding in first and second person forms in Leti and Moa; there are scattered citations from Roma and Luang-Sermata. It is regular in Wetan, where all cases involve *i* only, due to regular shift of *\*-u* > *i*. It is not found in Kisar (Rinnooy 1886, *passim*). Variation in all the Letic data suggests that binding may be (is becoming?) optional; Wetan also shows apparent extension of the rule by analogy, that is, *i* inserted in ‘wrong’ places, as in ‘correct’ *ira/dena* ‘they stay’ beside *ira/diena* ‘they stay’.

## 8.2 COMPOUNDING

(a) Not found in Fordata or Kei. In Fordata, most compounds simply drop the final V of the first member; Kei shows simple conjoining. Loss of binding in all the non-verb environments could be a F-K innovation, or independent in each language. Final-V loss in Kei, in any case, makes binding impossible.

(b) Yamdena: present but of uncertain productivity. A few cases are clear – *bali* ‘side’ + *duwe* ‘two’ > *bal’dyu* ‘both sides’. Others, such as the many tree names, may be frozen forms (e.g. *kabyarat*, *kafyalu*); as expected, no binding before #CC- (e.g. *kai-plasan*). A free form *kai* ‘tree’ is not listed. Other apparent compounds also contain undefined elements – *balnyote* ‘echo’ presumably < *bali* ‘answer’ + ?. Still other forms, as in Fordata, show loss of the final V.

(c) Selaru: present and apparently productive. In *r/sus twahar/ke* ‘they arrange thatching’ we see it within the verb phrase (*susu + tahar*).

(d) Lettic: present and productive, but variable, as Lettic *pipdiuma* ~ *pipi-duma* ‘sheep’ (< Lettic *pipi* ‘goat’ + Malay *domba* ‘sheep’). Very regular in Wetan; not found in Kisar.

## 8.3 SUFFIXING

(a) Absent in Fordata and Kei.

(b) Present in Yamdena, involving the plural marker: *lutur* ‘wall’, pl. *lutrwar*; *manik* ‘bird’, pl. *mankyar*. Inalienably possessed nouns behave somewhat differently.

(c) Present in Selaru, involving the singular marker *-ke*, as *hah* (Coward *hahy*) ‘pig’, *hahkye* (pl. *hahire*); in inalienables, cf. *lurik-kwe* ‘my bone’ (*luri+ku+ke*). The affected suffixes are first and second person possessives. There is a sizeable minority of forms which add *-ke* without binding.

(d) Lettic: The changes seen in Leti/Moa base forms and possessives are not examples of binding, but suggest a possible origin. Suppose that in Yamdena, final C had (or even now has, subphonemically) an echo-vowel, like Leti/Moa – thus, Yamdena *lutur* = [lúturu], *lutur+ar* > *lúturãr* > *lút(r)uar* = ‘*lutrwar*’. Such an explanation, however, is not so easily applied to Selaru.

Wetan possessives and plurals show V-deletion but no binding. Binding does occur in some phrases (perhaps better subsumed under ‘Compounding’, section 8.2), as in *nmati loe* ~ *nmatlyoe* ‘he died there’, or *ima ma* (*sic* for *imi ma?*) ~ *immi* ‘you all and...’. Josselin de Jong (1987:159) found this ‘*sandhi*’ to be dependent on the speed of speech.

## 8.4 FOSSILISED FORMS (?)

A number of Fordata, Yamdena and Selaru forms are cited with bound *y/w*. Drabbe suggests these are due to ‘alternate’ forms of some prefixes (e.g. *fa-* ~ *fau-*); perhaps too, some are due to loss of the first member of an original compound. A few examples: F *kyavu* ‘dust’; Y *nan/bwar* ‘glow (of fire, iron)’ cf. *au bware* ~ *au bare* ‘ember’; SLAR *hwalu* ‘a dove’ likely < \*\**manu+halu*. In Lettic all such examples involve *-i-*, reflecting prefixes, some known, as in Lettic *tialla* ‘road’ (locative \**i-* + \**talán* < \**Zalan*) and *siona* ‘digging stick’ (instrumental \**i-* + \**son* < \**suqan*), others uncertain as *kdioa* ‘twin’ (?\**ki-* + \**do* < \**nDuSa*) and *kdieli* ‘ring’ (?\**ki-* + \*\**deli*).

## 8.5 EFFECTS ON VERBAL CONJUGATION

Both Stresemann and Collins have reconstructed an 'Ambonese' (perhaps = CMP) verbal conjugation of Prefix + Verb. Using *\*taŋis* 'weep' as a model, we have:

SG	1. <i>*ku-taŋis</i>	2. <i>mu-taŋis</i>	3. <i>na-taŋis</i>
PL	1. <i>ta-taŋis, mi-taŋis</i>	2. <i>mi-taŋis</i>	3. <i>ra-taŋis</i>

Note that this structure is essentially the underlying conjugation (allowing for sound changes) of Fordata-Kei-Yamdena, Selaru and Lettic; it is also close to the surface forms of, for example, Tettum and Galoli (not to mention many other AN languages). Ambonese languages deleted the vowel of the prefix; nasals then assimilated to stem-initial C and went on to develop the same as original \*NC (see Stresemann 1927:119ff.); thus:

SG	1. <i>*ktaŋi</i>	2. <i>ndaŋi</i>	3. <i>ndaŋi</i>
PL	1. <i>ttaŋi, ndaŋi</i>	2. <i>ndaŋi</i>	3. <i>rtaŋi</i>

With later sound changes (e.g. *\*kt-* > *\*\*t*, *\*nd-* > *\*\*r*), the result was consonantal alternations in second and third person singular, and in first and second person plural, entailing changes in underlying forms and probable loss of some information.

Binding is at once innovative and conservative: innovative as to the sequence of rules that produce it (assimilation, deletion – which most likely began as fast-speech phenomena), yet conservative in that underlying forms are not changed, and little or no information is lost.

Both processes introduce asymmetries into the verbal paradigm. The Ambonese set (second and third person singular, first and second person plural) must have seemed particularly unmotivated, and it is no surprise that many of those languages have eliminated the irregularities by generalising one or another of the alternating consonants. The Yamdena-Selaru-Letic procedure (reanalysable as a simple insert *-u-* in first and second singular, and *-i-* in first and second plural) may in some sense be more logical, which perhaps accounts for its better preservation.

Since all these languages show verbal binding (even if only a remnant, as in Kei), it is assumed to have been present in their proto-language; very likely compound-binding was also present. It is hard to imagine, moreover, that morphophonemics of this complexity could be borrowed.

Selaru's *-ke* 'singular marker' is a notable innovation. As to the other environments, it would be tempting to view Yamdena, so conservative in other respects, as a model for the original situation, with individual losses and changes elsewhere, but we cannot be certain.

## 8.6 WIDER DISTRIBUTION OF BINDING

I recall encountering binding in Biak many years ago, before my interest had turned to the Lettic and Tanimbar languages, but the data were no longer to hand. My vague memories did not, I felt, warrant further discussion at this time of possible wider distribution, but Prof. Blust has brought to my attention the material in Dyen (1978, appendix), and the question must now be addressed.

Dyen (1978:248-249) cites evidence from Numfor (closely related to Biak), Patani (South Halmahera) and Kabhubhaka (West Sumba), which clearly shows 'verbal binding' in these three languages, quite comparable to that seen in Lettic and Tanimbar. In addition there is a West Sumba example involving the personal prefix *\*i* (*byuulu* < *\*i+buulu* 'pers. name') and two clear cases in the



Numfor numeral system. It is impossible to judge how extensive or productive the process may be in Patani or Kabhubhaka, whose examples could well be fossils (consider Kei's four verbs); in Numfor/Biak it appears at least as productive as in Fordata. Incidentally, Steinhauer (1985:471 ff.) gives the necessary documentation for binding in Biak.

I might also cite two suggestive Misool examples from Blust (1978:202-203): *syop* = Buli *sisop* 'bathe' and *falyan* = Buli *failan* 'mast'; Misool is said to belong to the South Halmahera West New Guinea (SHWNG) group. Binding is definitely absent from Buli, though syncopation of pretonic vowels is frequent there. It appears highly unlikely, I fear, that we have anything like a complete inventory of languages with binding – a major problem, of course, in the discussion that follows.

As Dyen (1978:248) says, binding "constitute[s] a reasonably strong argument for subgrouping together the languages exhibiting it", and I make just such use of it in positing the 'SE Maluku' subgroup. The need to explain the presence of binding in two additional widely separated areas would appear to weaken my argument, but does not necessarily do so.

If we proceed for now under the assumption that all languages with binding subgroup together, two hypotheses can be ignored, though not rejected out of hand:

(a) *Borrowing* from whatever direction, including non-AN source(s). This not only begs the question, but also violates one of historical linguistics' near axioms, to wit, that morphophonemic processes are hardly ever borrowed.

(b) *Parallel development* – but if, as I have suggested, binding originated in fast-speech rules, it may well be that details yet to be discovered could favour this possibility.

It is not utterly impossible that the SHWNG group, or some other group containing at least Numfor/Biak and Patani, could be an early offshoot of the 'Post-CMP' stage outlined in Table 9. None of the sound changes detailed in Blust (1978:192-193) is inconsistent with such a view, despite some local oddities (e.g. AN \*tTC > Buli *c*, Numfor *k*). We could then posit binding as an innovation in 'Post-CMP' which was totally lost in Proto Timor (if it is correctly subgrouped here), retained and perhaps extended in Proto SE Maluku, but only selectively retained in the Halmahera-New Guinea area. Perhaps there, as in Kei, final-V loss led to its restriction to word-initial environments.

The West Sumba case is more difficult, for these languages are very poorly documented: except for Wielenga's 1917 word list, where 'Kabhubhaka' is not mentioned (Dyen's *repe* 'hear' possibly = Wielenga's Laura dialect *rèngè* under the word *hooren*, no. 547). Wielenga speculates, from cultural evidence, that the languages of West Sumba may have a different origin from those of Eastern Sumba. The data in his word list, in my opinion, do not bear this out, as all Sumbanese languages appear to share a common ancestor. And even Wielenga points out that West Sumba has been isolated from the major trading centres and horse culture to the east. Interestingly, the languages of Sumba bear more resemblance to Savu (and so Roti?) than to languages of Flores (for instance, Sumba has merged the voiced and voiceless NC).

The weakest hypothesis here is:

(a) In West Sumba there is at least one area settled, somewhere in the past, by migrants who spoke a SE Maluku language with binding.



Slightly less weak would be:

(b) At least one language of West Sumba is the remnant of a formerly more widespread SE Maluku-type language with binding. The languages of East Sumba are interlopers, as also would be the main languages of Timor, Roti and Savu. I am uncomfortable with this because, aside from Dyen's two words, there is absolutely no other evidence for binding languages west of the Leti Archipelago.

The most attractive hypothesis would be:

(c) All languages of Sumba, along with Proto Lettic, Proto Tanimbar-Kei, Numfor-Patani (and perhaps Proto Timor and Savu) ultimately descend from the same 'Post-CMP' stage; binding has been lost in most of Sumba and SHWNG (and all of Timor/Savu, if included).

Thus we can see that it is not too difficult to create a reasonably coherent CMP subgroup, with binding, that embraces all the known descendants. The result, however, is to negate Blust's (1978) argument that SHWNG and Oceanic are coordinate under a node labelled 'Eastern Malayo-Polynesian'. Perhaps the situation could be salvaged by reformulating the SHWNG group without Numfor/Biak and Patani, though I suspect not. Blust's arguments for this subgroup seem very strong, however one may feel about its alignment with Oceanic.

More likely we should have to abandon our 'strong assumption' of a single subgroup with binding, and allow after all for parallel development within a relatively restricted geographic area in north-eastern Indonesia.

But if we reject parallel development, insisting on a single subgroup, and yet wish to retain Blust's EMP hypothesis, then we are forced into one of these uncomfortable positions:

(a) PTK/PLet/WSumba as well as SHWNG, are *all* coordinate with OC, and binding is an innovation within EMP; or

(b) PTK/etc. are indeed CMP languages and SHWNG is indeed EMP. Therefore, binding in both is *retained* from a higher level (Blust's East-Central MP at least) and has been lost in a very large number of daughter languages (most CMP and apparently all OC).

I do not know the answer.

## 9. WIDER RELATIONSHIPS

### 9.1. PREVIOUS ATTEMPTS

Only three attempts at classifying the languages of south-eastern Indonesia are known to me, all flawed in one way or another.

(a) S. J. Esser, on his well-known language map in the *Atlas van Tropisch Nederland*, posited a Timor-Ambon group, but without any further subdivision and without any record of his evidence. We can be sure, of course, that he was fully conversant with all published materials of the time, and probably had access to manuscript material and personal contacts with coworkers in the area. His published works on other languages demonstrate linguistic talent and insight of the highest order, so that even if his 'Timor-Ambon' group was impressionistically or intuitively based, it represents a very well-educated guess.

(b) Dyen (1965), in my view, gives short shrift to the Moluccan area. The manifold problems of his work are understandable: old and haphazard data, often amateurish and difficult to interpret; reliance on dictionaries; 'subadequate' lists; and extremely conservative decisions as to cognacy. The omission of Roti, one of the best described languages in all Indonesia, is inexplicable and a major flaw. The resultant 'Moluccan Linkage' is no more illuminating than Esser's grouping. And of course, the validity of lexicostatistics itself remains debatable, so that even if all the percentages are reliable (and they are not), the method offers little more than guidelines for further research and becomes irrelevant once enough data are available for proper comparison.

(c) It is wrong to call Collins (1982) 'flawed'; it is merely too brief vis-a-vis my area of interest. A few details aside, in fact, Collins reaches the same conclusion regarding Fordata-Kei-Yamdena as I do, and correctly notes the distinct position of Selaru. He has unearthed a new Tanimbar language, Selwasa, not mentioned by Drabbe. His comments on Teun, Nila and Serua are another first; these are apparently Lettic languages. He concludes:

...[T]he geographic boundary which separates Serua and Banda parallels a linguistic boundary. North of that line lie the numerous languages of Central Maluku; south of it are those languages most closely related to those studied by Jonker... (1982:124)

This boundary almost certainly extends eastwards to include at least Tanimbar-Kei; whether it includes Aru is not clear on the basis of available data.

## 9.2 SOUTHEAST MALUKU

I shall concern myself for now with justifying the alignment of Lettic, Selaru and PTK, mentioned already in section 7 and section 8 above. Space will permit only a brief glance in the direction of Roti, Timorese, Tettum and Galoli.

### 9.2.1 LETTIC

Based on Jonker (1932), Rinnooy (1886) and Josselin de Jong (1987), I have prepared a reconstruction of Proto-Letic (Mills:n.d.). All of these languages (Leti, Moe, Roma, Luang-Sermata, Wetan and Kisar) share not only much vocabulary but also these phonological features:

(a) changes in the vowel sequences produced by early loss of AN/CMP laryngeals. Ordered prior to loss of \*p;

(b) loss of AN \*p; shift of AN \*b > w (\*v);

(c) merger of AN \*d/D, r, j, R > r;

(d) merger of AN \*Z, t > t;

(e) shift of CMP \*mb, ŋg > p, k; \*nd > d;

(f) AN \*-uy > i, but \*-aw > a; distinct reflex of \*-ay;

(g) merger of AN \*n, ñ > n;

(h) binding: see section 8 above (presumed lost in Kisar);

(i) retention of all possible final C by addition of an echo-vowel; along with frequent syncope of the original ultima V, this produces the 'pseudometathesis' discussed in Mills and Grima (1980), for example, LET *ulti* ~ *ulit* 'skin' < \**kulit*, or KIS *hoplo* ~ *hopolo* 'to sail' < \*\**sombal*. Occurrence of

these forms is most likely governed by syntactic factors, though speech tempo may be a factor in some cases (Rinnooy's data are especially variable).

Features (a)-(h) are reconstructible for Proto Lettic; the question whether Proto Lettic had echo-vowels need not concern us here. The Proto Lettic sound system is set forth in Table 5.

TABLE 5: PROTO LETTIC SOUND SYSTEM

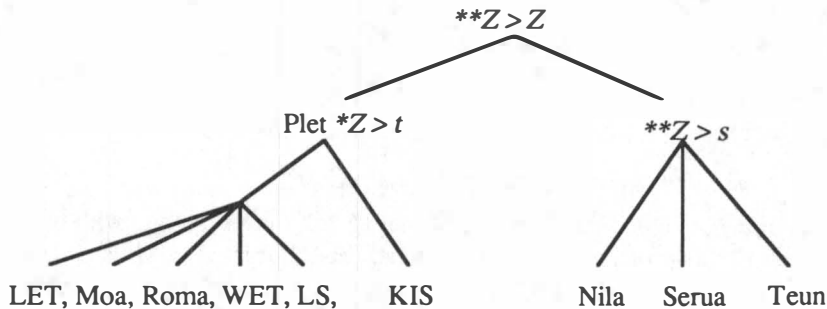
Consonants:	$p < **mb$	$t < *t, Z$	$k < **\eta g, g$	$? < *k$
	$v < *b$	$d < **nd, ns, nZ?$		
	$m$	$n < *n, \tilde{n}$	$\eta < *g$	
	$w < *w-w-$	$r < *r, j, R, d/D$	$l$	$y < *?-y-$
		$s$		
	$(\emptyset < *p, -w; \text{ for } *y \text{ see below})$			
Vowels:	$i < *i, -uy, -iw$		$u < *u$	
	$e < \emptyset_2, i(a), -ay, (i)$		$o < *u(a), (u, ??)$	
		$a < *a, \emptyset_1, -aw$		
Vowel sequences:		$ei < *a(i)$	$ou < *a(u)$	

In the modern languages, it is worth noting only these developments:

- (a) PLet  $*? > \emptyset$  everywhere except Moa  $-?$ , and questionably Roma  $h$  preceding a consonant.
- (b) PLet  $*\eta$  is reflected only in the correspondence KIS  $-m-$ : LET  $-kr- < *\eta-r\#$  in a handful of examples; otherwise  $*\eta > n$  everywhere.
- (c) AN  $*-ay >$  PLet  $*-e >$  Roma  $-a$ , LET, Moa, KIS, WET  $-i$ ; no LS examples.
- (d)  $*ei$  and  $*ou$  are retained as sequences or  $> e, o$  depending on environment (language specific). Most striking is KIS  $oi < *ei$ , as in *koi* 'excrement'  $< *taqi$ , *in/hoi* 'who'  $< *sa(i)$ .

This last feature, plus others (PLet  $*k > \emptyset$ ,  $*t > k$ ,  $*s > h$ , lack/loss of binding) sets Kisar apart from the other five, which are not differentiated much, and probably represent the remains of a former dialect chain.

Both Jonker and Brandes (see his introductory remarks to Rinnooy) held that the languages of the 'Damar Islands' were also related to Lettic; Collins' two pages (1982:123-125) on Teun, Nila and Serua are the first published data from the area. They evidently have  $r < *r, j, R, d/D$ . He refers to "consistent metathesis of the final consonants" which I take to indicate forms like LET *ulti* ~ *ulit* 'skin'. Further, "certain verbs...indicate inflectional change by internal vowel alternation," from which I infer binding. Serua, however, is said to reflect  $*Z$  as  $s$ ; if that is also true of Nila and Teun, then these languages can only be coordinate with Proto Lettic, related at a higher level, thus:



Since this area is known as Kepulauan Barat Daya ('Southwest Islands'), this higher node can be labelled 'Proto Barat Daya' and its two branches Proto Lettic and Proto Damar.

### 9.2.2 SELARU

In section 7 we have already discussed this language and its divergences from its Tanimbar neighbours, to wit.  $\emptyset < *p, k$ ,  $h \sim f < *b$  (note Nila  $\emptyset \sim f$ ),  $-i < *-ay, uy$ , and  $s < *Z$ . If we assume that  $*p$ , both here and in Lettic, was lost via  $**f \sim \emptyset \sim h > \emptyset$ , and that the reflexes  $t$  and  $s < *Z$  more likely stem from a voiceless than a voiced affricate, then Proto Lettic-Selaru would have had the following sound system:

TABLE 6: PROTO BARAT DAYA-SELARU (PBDS)

Consonants:	$f$ or $\emptyset < *p$	$t$	$c < *Z$	$? < *k$
	$v < *b$			
	$b < **mb$	$d < **nd$		$g < **\eta g, g$
	$m$	$n$		$\eta$
	$w$	$r < *r, j, R, d/D$	$l$	$y$
		$s$		
Vowels:	as in PLet above, except $*ua$ , $*ia$ remain as sequences.			

### 9.2.3 PROTO SE MALUKU

At this point we can compare PTK, and note that only three mergers and a few phonetic shifts distinguish PTK and PBDS. The mergers involve  $*d/D$  and  $*Z$ , the vowel sequences  $*ei$  and  $*ou$ , and the diphthongs  $*-ay$  and  $*-uy$ ; the phonetic shifts involve  $*p$ ,  $*b$  and the  $*NC$ . We can therefore unite PTK and PBDS under 'Proto SE Maluku' with a sound system as set forth in Table 7.

We have seen that PTK merged  $*d, Z > *d$ , while PBDS merged  $*d, r > *r$  and presumably devoiced  $*Z$  to  $**c$ ; PTK lost  $*-y$  while PBDS changed vowel quality; PTK simplified  $*ei$ ,  $ou$  to  $*e, o$ , PBDS retained them.

TABLE 7: PROTO SE MALUKU

Consonants:	$f-f-p$	$t$	$k$
	$b$	$d < *dD$	$Z$
	$mb$	$nd$	$\eta g < *\eta g, g$
	$m$	$n$	$\eta$
	$w$	$r < *r, j, R$	$l$ $y$
Vowels:	$i < *i, -iw$		$u < *u$
	$e < *e_2, (i)$		$o < (u, \text{loans?})$
		$a < *a, e_1, -aw$	
Diphthongs:	$-ay$	$-uy$	
Sequences:	$-ei- < *a(i)$	$-ou- < *a(u)$	

As shown in section 8, binding affects all the languages of this subgroup, and is viewed as a distinctive innovation.

## 9.3 DESCENT FROM PROTO CENTRAL MALAYO-POLYNESIAN

Let us now compare Proto SE Maluku with the probable Central Malayo-Polynesian proto-language, assuming its sound system as in Table 8.

TABLE 8: PROTO CENTRAL MALAYO-POLYNESIAN (?)

Consonants:	<i>p</i>	<i>t</i>		<i>k</i>	<i>q</i>	Vowels: <i>i ə a u</i>
	<i>b</i>	<i>d</i>	<i>Z</i>	<i>g</i>		
	<i>mb</i>	<i>nd</i>		<i>ŋg</i>		Diphthongs: <i>-ay, -aw, -uy</i>
	<i>m</i>	<i>n</i>		<i>ŋ</i>		
	<i>w</i>	<i>r,l</i>	<i>y</i>	<i>R</i>	<i>S</i>	
		<i>s</i>	<i>j</i>			

NOTE: This assumes, crucially, the merger of AN *\*d* and *\*D* > CMP *\*d*, and merger of the voiced and voiceless nasal-grade consonants (though likely not all the dental/palatal NC had merged). Reduction of the number of laryngeals, merger of *\*t,T,C* > *\*t*, *\*s,c* > *\*s*, *\*n,ñ* > *\*n* and *\*iw* > *\*i* are also assumed to have taken place in pre-CMP times, but even if they had not, they are of slight diagnostic value in eastern Indonesian languages.

For now, note that PSEM shows a) loss of *\*-w*, b) merger of *\*r,j,R*, c) loss of *\*q,S* and creation of the new sequences *\*\*e-i-* and *\*\*ou--*, d) *\*ə* > *\*\*e-a-*.

## 9.4 LANGUAGES OF TIMOR

If we are on anything like firm ground in positing common ancestry for PTK, PLet and Selaru, the footing becomes slippery indeed as we approach Timor. The island's principal languages – Tettum, Galoli, Timorese (Dawan) – are adequately described, even exhaustively so in the case of Roti, which is not, of course, spoken on Timor but appears to subgroup better with its languages than with any others. Capell's (1944) survey offers bits of data for other Timorese languages, including some non-AN ones, but his 'Vaikenu' is clearly a dialect of Timorese, 'Belu' probably a dialect of Tettum. The others cannot be classified at this point.

## 9.4.1 CHANGES IN COMMON

Aside from Tettum and Galoli, there is little obvious evidence of a close relationship between these four languages. There are enough shared developments, however, that a case for a Timorese subgroup can be made.

- (a) All four show  $\emptyset$  < AN *\*R*.
- (b) AN diphthongs: *\*-ay* > *e*, *\*-aw* > *o*, *\*uy* > *i* in all.
- (c) AN *\*q,S* >  $\emptyset$ ; *V(qS\emptyset)V* sequences > *\*a()i* > *e*, *\*a()u* > *o* in all.

(d) Following from (b) and (c), all four languages *probably* have 7-vowel systems *i e ε a u o ə* (*ε* and *ə* hereafter 'è' and 'ò'). The evidence for contrastive mid/low vowels is ambiguous: all the writers describe such contrasts and indicate them orthographically, but also cite 'rules' for determining where the vowels occur.

(e) Laxing/devoicing of AN *\*p* and *\*b*, likely via *\*\*ϕ* and *\*\*v*: *\*\*ϕ* > ROT *h-* and *-ϕ-* or *-ʔ-*, TIM *h-* and *-h-* or *-ϕ-*, TET *h-ϕ-s*, GAL *ϕ-ϕ-s*; *\*\*v* > ROT, TIM *f-f-*, TET *f-h-*, GAL *h-h-*.

- (f) AN  $*n, \tilde{n}, \eta > n$  in all.
- (g) None of the languages has binding; conjugation is by prefix + verb.
- (h) None of the languages shows  $y-$ .

#### 9.4.2 DISTINCTIONS

To my view, these are at least as weighty as the shared sound changes, involving as they do the crucial elements  $*j$ ,  $*Z$ , and  $*d/D$ , along with  $*ə$  and the  $*NC$ .

(a) Roti merges  $*j, Z > d$ , and  $*dD, r > r \sim l$  depending on dialect. The  $*NC$  appear as /NC/, voiced /C/ or /N/ depending on position and dialect. AN  $*ə > e$  in both penult and ultima – the latter in contrast to all the languages discussed in this paper.

(b) Timorese merges  $*j, dD, r, l, n > n$ , probably ordered 1)  $*j > d$ , 2)  $*d > r$ , 3)  $*r > l$ , 4)  $*l > n$ . AN  $*Z$  (via  $**d?$ )  $> r \sim l$  depending on dialect (there is also a final  $-l$  of unclear origin).  $*ə > e-a-$ ;  $*mb, \eta g > p, k$ ;  $*nd > k$ , perhaps dial.  $r$ , as Middelkoop cites  $k \sim r$  variants.

(c) Tettum shifts  $*ə > o-a-$  and  $*Z > d$ ; Galoli has  $*ə > e-a-$  and  $*Z > s$ . Otherwise they share  $*dD, r > r$ ,  $*nd > d$ ;  $*mb, \eta g > TET GAL b, k$ , though GAL also has  $p, g$  reflexes. All TET/GAL examples for  $*j$  present problems, but  $r$  is the most likely reflex (the alternative is  $l$ ).

It is theoretically possible, of course, to create a proto-language for any group of ultimately related languages (e.g. 'Proto Swedish-Hindi!'), but such a creation will have little value except as an intriguing hypothesis, or, to put it more bluntly, linguistic game-playing. In that spirit, I offer the following tentative thoughts on 'Proto Timor' and its possible connection with Proto SE Maluku (PTK-PBD-SLAR).

The developments of  $*j$ ,  $*dD$ ,  $*Z$  and  $*ə$  are not as confused as they appear.  $*Z$  and  $*ə$  obviously survive well down into the family tree;  $*dD$  evidently  $> **r$  very early on, after which it can be posited that  $*j > **d$ ; many of the shared changes outlined in section 9.4.1 also reflect early change. Given four languages, multiple alignments are possible: the three  $e < *ə$  languages versus interloper Tettum with  $o$ ; or, the three 'mainland' languages versus Roti, among others. But there is sufficient other evidence (from vocabulary and grammar) that Tettum and Galoli do group together; if that assumption is made, the question then is, what of Roti and Timorese?

Roti and Timorese share a development in common that may be relevant: at some point in their history, both languages lost all final consonants, with subsequent replacement. In Roti, the replacements included the possessive markers ( $-k$ ,  $-m$ ,  $-n$  etc.) on inalienable nouns, a 'verbal'  $-k$ , 'nominal'  $-k$  and  $-s$ . Presumably there were once transparent alternations between CVCV  $\sim$  CVCV+C forms, but nowadays, all but the possessives seem fossilised, and the +C forms often show semantic changes.

Timorese shows similar developments, but has additional finals, some of which are quite productive:  $-f$  creates 'generics' out of inalienables, for example, *atèk* 'my liver', *atèf* 'liver in general'; or  $-b$  'causative/transitive'; or  $-s$ ,  $-t$ ,  $-k$  for verbal or nominal derivatives. Many plant names have  $-l$ . Not only are the origins of these very obscure; they contribute to the overall impression that Timorese has undergone influences unseen anywhere else in the area – in particular, the metathesis of final -CV, evidently syntactically conditioned (cf. *teun* 'three', *katenu* 'third'), which is so reminiscent of that seen in Rotuman.



The vowel system underlying all four languages must have had *\*i*, *\*u*, *\*a* and *\*ə* (retained < CMP), with *\*\*i* < *\*-uy*; *\*\*e* < *\*-ay* and *\*\*o* < *\*-aw*; and likely *\*\*e* < *?\*ei* < *\*aθi* and *\*\*o* < *?\*ou* < *\*aθu*. Since 'diphthongal' *-e* and *-o* only occurred unstressed and finally, perhaps they were [-tense] *e:ə*, *o:ə* whereas 'sequential' *-e-* and *-o-* were stressed, in #C\_\_(C)# position, and were possibly contrastive *e:ə*, *o:ə* depending on open versus closed syllables. Perhaps too, *e* *o* were present in Ambonese and outside loans. These intriguing ideas, I must confess, have yet to be researched fully, but they might account for the seven-vowel system *i e ə a u o ə* which, after shift of *\*ə* > *e/ə* or *o/ə* (similarly conditioned?), these languages appear to share.

From all the foregoing, I posit 'Proto Timor', with two branches: West Timor (Roti, Timorese) and East Timor (Tettum, Galoli). But I must re-emphasise the tentative nature of the alignment, with its totally phonological basis, unlike 'Proto SE Maluku' (PTK-PBD-SLAR) where binding, a unique and complex phenomenon, offers an additional and convincing basis for subgrouping.

### 9.5 PROTO SOUTHERN MALUKU

It is now possible to conceive of a 'Post-CMP' stage from which both Proto Timor and Proto SE Maluku can derive. Minimally, this stage must show CMP *\*q, S* >  $\emptyset$  with *\*a()i* > *\*\*ei-* and *\*a()u* > *\*\*ou-*, but CMP *\*j*, *\*R*, *\*r*, the three diphthongs and *\*ə* remain distinct. Perhaps *\*\*e* and *\*\*o* are also present due to borrowing plus sporadic lowering of *\*i* and *\*u*. Only minor contortions are needed to derive PSEM; the number of changes required for Proto Timor suggests rather strongly that additional intervening stages are probably needed here. If not, however, then this 'Post-CMP' stage can be labelled 'Proto Southern Maluku', as is depicted in Table 9.

Viewed against other languages of the area, 'Proto Southern Maluku' is at least distinguished by loss of the laryngeals and resultant vowel changes. A laryngeal reflex and the vowel sequences remain in Bima and Manggarai, for instance; Savu loses the laryngeals but retains the vowel qualities. Bima and Manggarai, further, even show *apparent* reflexes of both voiced and voiceless NC, and on that basis cannot derive from CMP as it is presently conceived. Yet without the NC merger, CMP almost ceases to be a unique entity. Clearly, much interesting work is yet to be done in the Lesser Sunda/Moluccan area.

TABLE 9: FROM 'POST-CMP' TO THE TIMORESE AREA

CMP (See Table 8)

\*p > f/\_ V; \*g,ŋg > \*ŋg; \*q,S > Ø: \*a(i) > \*\*ei, \*a(u) > \*\*ou

'Proto Post-CMP'

Consonants:	f-f-p	t		k	Vowels:	i		u
	b	d	Z			e?	ə	o?
	mb	nd		ŋg			a	
	m	n		ŋ		-ei-		-ou-
	w	r l	y	R				
		s	j			-ay	-aw	-uy

PROTO TIMOR

\*R > Ø; \*-p > f: all \*f > φ(?);  
 \*b > v; \*d > r: \*j > d; \*ŋ > n;  
 \*y- > Ø; \*ay > è, \*aw > ò, \*uy > i;  
 \*k begins? > ?; \*ei > e/è?, \*ou > o/ò?

PROTO SOUTHERN MALUKU

\*j,R,r > r; \*-w > Ø;  
 \*ə > e-a-

f or φ	t		k, ?	i		u
v	d	Z		e?	ə	o?
mb	nd		ŋg	-è	a	-ò
m	n			ei or e/è		
w	r l	-y-		ou or o/ò		
	s					

PROTO BARAT  
 DAYA-SELARU  
 (see section 9.2 above)

PROTO  
 TANIMBAR-KEI

PROTO WEST TIMOR

-C > Ø: replaced; \*ə<sub>2</sub> > e/è  
 \*f/φ > h; \*v > f; 8 vowels

PROTO EAST TIMOR

\*-f/φ > s; \*ə<sub>1</sub> > a; \*d > r(l?)  
 \*f/φ > h; \*v > f; 8 vowels  
 \*mb, nd, ŋg > b, d, g

ROTI

\*h > h-?  
 \*r > r ~ l  
 \*Z, d > d  
 \*ə<sub>1</sub> > e  
 \*NC > NC ~ C ~ N

TIMORESE

\*h > Ø/V\_V  
 \*d > r > l > n  
 \*Z > (d?) > r ~ l  
 \*ə<sub>1</sub> > a  
 \*mb, nd, ŋg > p, k, k

TETTUM

\*h > Ø/V\_V  
 \*f > h/V\_V  
 \*Z > d  
 \*ə<sub>2</sub> > o  
 \*g > k

GALOLI

\*h > Ø  
 \*f > h  
 \*Z > (c?) > s  
 \*ə<sub>2</sub> > e  
 \*b, g > b ~ p, g ~ k

## BIBLIOGRAPHY

## ABBREVIATIONS

<i>BKI</i>	<i>Bijdragen tot de Taal-, Land-, en Volkenkunde</i>
<i>TBG</i>	<i>Tijdschrift voor Indische Taal-, Land- en Volkenkunde</i>
<i>VBG</i>	<i>Verhandelingen van het Batavi-aasch Genootschap van Kunsten en Wetenschappen</i>
<i>VKI</i>	<i>Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde</i>

- BLUST, R. A.**, 1978, Eastern Malayo-Polynesian: a subgrouping argument. In Wurm and Carrington, eds 1978:181-234.
- 1981, The Soboyo reflexes of Proto-Austronesian \*s. In R.A. Blust, ed. *Historical linguistics in Indonesia*, part I, 21-30. *NUSA* 10. Jakarta: Universitas Atma Jaya.
- CALON, L.F.**, 1890, Woordenlijstje van het dialect van Sikka. *TBG* 33:501-530; 34:283-363.
- CAPELL, A.**, 1944, Peoples and languages of Timor. *Oceania* 14/3:191-219, 14/4:311-337, 15/1:19-48.
- 1968, *A new Fijian dictionary*. 3rd edn; 1st, Sydney, 1941. Suva, Government Printer/London: Eyre & Spottiswoode.
- COLLINS, J. T.**, 1980, The historical relationships of the languages of Central Maluku, Indonesia. PhD dissertation, University of Chicago (since published as *PL*, D-47, 1983).
- 1981, Preliminary notes on Proto-West Central Maluku: Buru, Sula, Taliabo, and Ambelau. *NUSA* 10:31-45.
- 1982, Linguistic research in Maluku: a report of recent field work. *Oceanic Linguistics* 21/2:73-146.
- COWARD, N.E.**, 1989, A phonological sketch of the Selaru language. In: Wyn Laidig, ed. *Workpapers in Indonesian languages and cultures*, vol.7:1-32. *Maluku*. Ambon, Pattimura University/SIL.
- DEMPWOLFF, O.**, 1934-37-38, *Vergleichende Lautlehre des austronesisches Wortschatzes. Zeitschrift für Eingeborenen Sprachen*, Beihefte 15, 17, 19.
- DORES, R. das**, 1907, *Diccionario Teto-Português*. Lisboa: Imprensa Nacional.
- DRABBE, P.**, 1926a, *Spraakkunst der Fordaatsche taal*. *VBG* 67/1.
- 1926b, *Spraakkunst der Jamdeenske taal*. *VBG* 67/2.
- 1932a, *Beknopte spraakkunst en korte woordenlijst der Slaroeësche taal*. *VBG* 71/4.
- 1932b, *Woordenboek der Fordaatsche taal*. *VBG* 71/2.
- 1932c, *Woordenboek der Jamdeenske taal*. *VBG* 71/3.
- DYEN, I.**, 1965, *A lexicostatistical classification of the Austronesian languages*. *IJAL* Memoir 19. Baltimore: Waverly Press.
- 1978, The position of the languages of eastern Indonesia. In Wurm and Carrington, eds, 1978:235-254.
- EARL, G. W.**, 1848, Specimens of the dialects of Timor and the chain of islands extending thence to New Guinea. *Journal of the Indian Archipelago and Eastern Asia* 2:695-704.
- EIJBERGEN, H. C. van**, 1864, Korte woordenlijst van de taal der Aroe- en Kei-eilanden. *TBG* 14:557-568.
- GEURTJENS, H.**, 1921a, *Spraakleer der Keieesche taal*. *VBG* 63/2.
- 1921b, *Woordenlijst der Keieesche taal*. *VBG* 63/3.
- JONKER, J. C. G.**, 1908, *Rottineesch-Hollandsch woordenboek*. *VBG* 48/1-2. Leiden, E. J. Brill.
- 1932, *Lettineesche taalstudiën*. *VBG* 69/7.
- JOSSELIN DE JONG, J. P. B. de**, 1987, *Wetan fieldnotes: some eastern Indonesia texts with linguistic notes and a vocabulary*. Dordrecht: Foris. *VKI* 130.
- MATHIJSEN, A.**, 1906, *Tettum-Hollandsch woordenlijst*. *VBG* 56/2.
- MIDDELKOOP, P.**, 1939, *Amarasisch-Timoreesche teksten*. *VBG* 74/2.
- 1950, *Proeve van een Timorese grammatica*. *BKI* 106:375-517.
- MILLS, R. F.**, n.d., *Reconstruction of Proto-Letic*. MS.

- MILLS, R.F. and J. GRIMA, 1980, Historical developments in Lettinese. In: Paz Buenaventura Naylor, ed. *Austronesian studies: papers from the Second Eastern Conference on Austronesian Languages*, 273-284. Ann Arbor: University of Michigan, Center for South and Southeast Asian Studies.
- PIETER, C. et al., 1986, *Struktur Bahasa Tanimbar*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa, Dep. Pendidikan dan Kebudayaan.
- RINNOOY, N., 1886, Maleisch-Kissersche woordenlijst. *TBG* 31:149-213.
- SILVA, M. M. A. da, 1905, *Diccionario Portugues-Galoli*. Macau.
- STEINHAUER, H., 1985, Number in Biak: counterevidence to two alleged language universals. *BKI* 141/4:462-485.
- STOKHOF, W.A.L., ed. ... with Lia SALEH-BRONKHORST and Alma E. ALMANAR, 1981, *Holle lists: vocabularies in languages of Indonesia*, vol. 3/1: *Southern Moluccas; Central Moluccas: Seram (1)*. (*MLIn* 4). *PL*, D-35.
- STRESEMANN, E., 1927, *Die Lauterscheinungen in den ambonischen Sprachen*. *Zeitschrift für Eingeborenen Sprachen*, Beiheft 10.
- TETELEPTA, J. et al., 1985, *Struktur Bahasa Kei*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa, Dep. Pendidikan dan Kebudayaan.
- WARD, J. H., 1969, Some background for a linguistic survey of the Indonesian Lesser Sunda Islands. *Working Papers in Linguistics, University of Hawaii* 1/4:169-212.
- WIELENGA, D.K., 1917, *Vergelijkende woordenlijst der verschillende dialecten op het eiland Soemba*. *VBG* 61/5.
- WURM, S.A. and Lois CARRINGTON, eds, 1978, *Second International Conference on Austronesian Linguistics: proceedings*. *PL*, C-61.



## THE FOSSILISED AFFIXES OF BIKOL

MALCOLM W. MINTZ

### 1. INTRODUCTION

This paper is a preliminary analysis of the fossilised affixes of Bikol based on data drawn from the Marcos de Lisboa *Vocabulario de la lengua Bicol*, compiled between 1609 and 1613. The spelling of all entries has been regularised and follows the orthography in the modern *Bikol-English dictionary* (Mintz and Britanico 1985).

What makes the Lisboa dictionary a fine source for identifying fossilised affixes is its rich store of vocabulary no longer found in the modern, standard Bikol based on the dialect of Naga City. Conclusions reached in this paper relate only to standard Bikol and not to other dialects of the Bikol region.

The paper attempts to identify earlier affixes which are not currently recognised as separate morphemes. The discovery techniques used are recurrence and contrast. Comparative evidence is presented where relevant.

Underlying the discussion is an assumption that the Bikol word base most commonly comprises two syllables and that words of more than two syllables have therefore been formed by processes such as affixation (the addition of prefixes, infixes and suffixes) and reduplication. Many of the morphemes identified as possible affixes may have served in creating disyllabic word bases from monosyllabic roots. This possibility is not pursued further in this paper. For a general discussion of Austronesian root theory the reader is referred to Blust (1988).

Word bases with an internal consonant cluster, (C)VCCV(C), were examined to see if these were composed of a prefix followed by a root of the more common (C)VCV(C) structure. While some evidence was found to support this, a convincing case could not be made. Onomatopoeic word bases – that is, word bases which Lisboa defines as describing sounds – have been excluded from the



discussion. The polysyllabic nature of such bases appears to derive from the need to approximate the sound described as closely as possible.<sup>1</sup>

The sections below are headed by the particular affix being investigated. Not all of these affixes, however, are eventually proved as independent morphemes.

## 2. THE PREFIX *a-*

Lisboa identifies at least three examples as having alternative *a-* affixed and unaffixed forms.

In example (1) Lisboa cites the pair *atitiras/titiras* and, although he makes no reference to *tiras* as a word base, the semantic relationship between this base and the affixed forms is obvious. It is not obvious why the affixed forms appear to involve reduplication of the first CV of the root since no other examples cited by Lisboa as equivalent involve such reduplication. It is probable that *a-* is not an independent affix here. The affix in question is possibly *ati-* with *ti-* remaining after loss of the initial vowel (see section 6).

In example (2) *ali-* may also be a fossilised prefix (see section 4) and *li-* the remaining form after the loss of the initial *a-*, but the Lisboa dictionary contains no evidence to support *púdok* as an independent root.

With example (3) Lisboa refers to *ha'ín* as an alternative form. In modern Bikol only *ha'ín* is used.

- |     |                         |  |
|-----|-------------------------|--|
| (1) | <i>atitiras/titiras</i> | caterpillar (type which burns like fire when touched); MAG- to eat something (of this type of caterpillar)           |
|     | <i>tiras</i>            | MA- to burn or scald the mouth when eating something hot or chewing something piquant (such as an areca nut mixture) |
| (2) | <i>alipúdok/lipúdok</i> | MAG- to be sad or melancholy   |
| (3) | <i>aha'ín/ha'ín</i>     | where (used only to inquire about the location of things present at the time the question is being asked)            |

In addition to those entries identified as equivalent by Lisboa there are other pairs which can also be identified.

- |     |                    |  |
|-----|--------------------|--|
| (4) | <i>agamad</i>      | MAG- to know how to reason something out; to know how to approach something wisely or prudently          |
|     | <i>gamad-gamad</i> | MAG-, -AN to fathom; to understand something in depth  |
| (5) | <i>apurong</i>     | plant (type which climbs like ivy, producing white, odourless flowers); MAG- to climb wildly (any plant) |

<sup>1</sup>Word bases presented as examples are also cited with their relevant affixes. For example, the word base *atitiras* or *titiras* in example (1) means 'caterpillar' (the specific species described). When affixed with MAG-, the infinitive forms being *magatitiras* or *magtitiras*, it has the verbal meaning 'to eat' (referring to the actions of this caterpillar). Where particular bases are used only when affixed, no independent meaning is given for the base. *tiras*, also found in example (1), has no independent meaning. When affixed with MA-, *matiras*, the meaning is verbal: 'to burn or scald the mouth...'. Many entries are cited with affix pairs. *adláng* in example (6) is cited with the affix pair MAG-, -ON. This is an abbreviated form for the verbal infinitives *magadláng* and *adlangón*, the first indicating an agentive sentence subject, and the second an accusative one.

	<i>purong</i>	well-twisted cord or rope; MAG-, -ON to produce well-twisted cord or rope
(6)	<i>adláng</i>	MAG-, -ON to scrape the surface of a type of gold called <i>minambóg</i> with an instrument called <i>tugtohon</i> to discover if it masks a higher quality gold below its surface
	<i>dulang</i>	MAG-, -ON to work the mines; to look for gold in rivers, beneath one's land, etc.; to work metal after it has been forged

One further pair mentioned as equivalent by Lisboa may help explain the origin of what appears to be the prefix *a-*. Neither of the forms mentioned by Lisboa in example (7) is the modern Bikol form *pa'nó*. The modern meaning has also changed and is more closely related to the English 'how' as in 'How do I cook rice?', although it may also be used in expressions related to those presented below, such as *Pa'nó kayá!* 'How do you think it happened!'.

(7)	<i>apa'nó/napa'nó</i>	how (used as expression: <i>Apa'nó nang ínit in!</i> How very hot this is!, <i>Apa'nó nang bansáy!</i> How beautiful!)
-----	-----------------------	--

*Pa'nó* is composed of the causative prefix *pa-* and the root *anó* 'what': *pa'anó*. The prefix *na-* is probably the stative verbal prefix which would normally have a long vowel, *na:-*, although length is not indicated by Lisboa and is not found in any of the modern Bikol orthographies. The full form is clearly *napa'anó*, which is reduced to *napa'nó*. There may have been a further reduction from *napa'nó* to *apa'nó*, this last form later being treated as a word base with no affixation.

Word bases such as those presented in examples (3)-(6) may first have been prefixed with *na:-* as part of the regular process of inflection. The initial *n* could have been lost, giving rise to doublets with and without an initial *a-*. The motivation for loss of the initial nasal, however, is not clear as there are no verbs prefixed only with *a-*, and the stative prefix *na:-* remains in use up to the present.

There is, however, a prefix of limited occurrence of the form *a-* in modern Bikol, which is also mentioned by Lisboa. This form is prefixed to non-subject agent pronouns and markers of the class represented by *ni*. A cognate with similar function is found in Tagalog in examples such as *ani Pédro* 'Pédro said' (Blake 1925:136). *niámó'* and *niáto'* in example (8) have only a literary use in modern Bikol. The prefix referred to appears consistently as the *a-* in *a- + ko'*, *a- + mo'* and *a- + to'*.

(8)	<i>niáko'/aniáko'</i>	I said
	<i>niámo':</i>	
	<i>haróng niámo'</i>	our (EXCL) house
	<i>sábi niámo'</i>	we (EXCL) said
	<i>niáto:</i>	
	<i>haróng niáto'</i>	our (INCL) house
	<i>sábi niáto'</i>	we (INCL) said

This prefix can also be found in the old Bikol conjunctions *asi* (*a- + si*) 'and (SG)' and *asa* (*a- + sa*) 'and (PL)'. While these are no longer used in modern Bikol, the current literary form *asin* 'and (SG and PL)' can probably be traced back to these for its origin. Example (9) shows both *asi* and *asa* used in context.

(9)	<i>sa Pédro asi Juan</i>	Pedro and Juan
	<i>sa Pédro asa Juan</i>	Pedro, Juan and those with them

While a case for a prefix of the form *a-* has been made, it is difficult to see how this prefix, restricted to pronominals and nominal particles, can be shown to be actively used on verbs. An explanation based on deletion of the initial nasal is more tenable.

A similar type of change based on nasal deletion may also explain the occurrence of the prefix *aN-* discussed below, although there the motivation for the change is clearer.

### 3. THE PREFIX *aN-*

While Lisboa identified only one pair where the alternative form is prefixed with *aN-*, presented as example (10), and one other related pair of words, presented as example (11), there are other entries showing *aN-* affixed and unaffixed pairs. The process of reduplication (of the first CV of the root prior to or following affixation) has also played a part in word formation.

- (10) *lana/anlalana* type of bird (small, with blue plumage, commonly found wading in shallow water in search of fish)

In addition to *amamaliw* in example (11), Lisboa cites *umamaliw* as an alternative form, probably the result of *-um-* being affixed to a stem of the form *amaliw* at a later stage in the derivation processes (see section 10).

- (11) *amamaliw* whale; large fish  
*baliw* MAG-, -ON to devour a person; to devour a boat with all its crew; to flood a village (of a large sea mammal or fish called *amamaliw*)

While Lisboa refers to no other related pairs of entries, other related forms can be identified.

- (12) *anunuktók* type of bird (small, slightly larger than a swallow)  
*tuktók* MAG- to peck (birds)

The two words in example (13) are not obviously related. In addition to the affixation of *aN-* the derivation involves the affixation of the plural infix *-Vr-*.

- (13) *andaramoy* type of tree (used for curing abscesses); a plaster made from the secondary bark is placed on the abscess, opening it and drawing out the pus so that it may heal  
*damoy* MAG-, -ON to wash off soot, grime etc. smeared on the body; MAG-, I- to wash such soot or grime with water; to smear with soot or grime; MANG- to wash the face or other part of the body; to wash someone else

With reference to example (14) the modern form is *alimpupúro* (see section 3).

- (14) *ampupúro* the crown of the head where the hair appears to spiral out in different directions; also refers to the part of an animal where the fur or hair spirals out in the same way  
*puró* the top or point of something  
(15) *amatong* a snare or trap for catching monkeys; MAG-, -ON to trap monkeys with an *amatong*

	<i>batong</i>	wide mesh net used for catching wild boar, deer and other game; MAG-, -ON to hunt for particular game with such nets
(16)	<i>amúkid</i>	wild abaca found in the mountains, not suitable for making cloth or rope
	<i>búkid</i>	hills, mountains

All of the forms affixed with *aN*- above appear visually similar to the general verbal affix *maN*- minus the initial nasal.

In the examples above, only *damoy* is shown as taking the *maN*- prefix. The affixed singular infinitive form appears as *mandamoy*, and the plural as *mandaramoy*. The future and present tense forms involve reduplication of the first CV of the root; and the past and present, the substitution of the *m* of the prefix by *n*. The future, present and past tenses of the plural form of *mandaramoy* are *mandadaramoy*, *nandadaramoy* and *nandaramoy* respectively. The nominal form under discussion, *andaramoy*, would result if the initial nasal of the infinitive or past tense forms were lost.

The tie between *damoy* and *andaramoy*, however, is not obvious since there seems to be little relationship between the tree with its poultice made from bark, and the act of washing. The fact that *damoy* is not a common word for wash and the fact that it is only used to wash off soot or grime smeared on the body, may indeed indicate a closer relationship than is apparent. In any case, nothing further can be proved from this pair of entries.

The prefix *maN*- is commonly used when a general search is made for the object indicated by the root, or at the location indicated by the root, or when using the object indicated by the root. Examine the following examples:

(17)	<i>abúkay</i>	type of parrot (large, white); MANG-, PANG-ON: <i>mangabúkay</i> , <i>pangabukáyon</i> to go in search of the <i>abúkay</i>
(18)	<i>babáyi</i>	woman; MANG-, PANG-ON: <i>mamabáyi</i> , <i>pamabayihon</i> to go in search of women; to associate with or be friends with women
(19)	<i>bakulod</i>	rocky shoals found in the sea; MANG-, PANG-ON: <i>mamukulod</i> , <i>pamukulodon</i> to search for shellfish in such shoals
(20)	<i>kalásag</i>	shield; MAG-, -ON to carry a shield; MAG-, -AN to protect with a shield; MANG-, PANG-ON: <i>mangalásag</i> , <i>pangalásagon</i> to fend off blows with a shield; to use a shield for protection in battle

It is possible to reconstruct forms such as the following for examples (10)-(16) by analogy with the forms presented in examples (17)-(20): *nanlalana* 'to go in search of the *lana*'; *namamaliw* 'to attack or devour (of a large sea mammal or fish)'; *nanunuktók* 'to peck, to go around pecking (birds)'; *nampupúro* 'to go to the top or point'; *mamatong* 'to hunt with the net called *batong*'; *mamúkid* 'to go to the mountains in search of something'. A subsequent change leading to the formation of a noun associated with the action of the verb, and the deletion of the initial nasal by analogy with the nominal particle *aN*, could then result in nominals which appear to be prefixed with *aN*-.

4. THE PREFIXES *ali(N)-*, *li-* AND *ari(N)-*4.1 THE PREFIX *ali(N)-*

The headword entries in the Lisboa dictionary present striking evidence for the existence of a prefix of the form *ali(N)-*, a prefix noted elsewhere for Austronesian languages (Blust 1980:24, 1983-84:35). A provable relationship on the basis of contrast between *ali(N)-* affixed and unaffixed forms can only be made for a small number of entries. The affix, however, does appear to recur in a large number of entries, some of which are presented below.

<i>alibagtá</i>	type of plant (wild, producing sprays of red flowers)
<i>alibangbáng</i>	butterfly
<i>alibutbót</i>	type of plant (producing a red fruit)
<i>aliwasog</i>	MAG- to stretch (one who is still sleeping, though about to wake up)
<i>alilintá</i>	type of leech (found in wet and swampy areas)
<i>alimato</i>	type of leech (found in the forests and mountains)
<i>alimbabayod</i>	millipede
<i>alimusog</i>	pulp or inner core (of the tusks of an elephant, of woody plants)
<i>alimputo</i>	the highest reaches of a tree
<i>alintatáw</i>	pupil of the eye
<i>alintuhod</i>	type of plant (which in times of famine is cooked mixed with rice)
<i>alitagtág</i>	type of tree (possessing a hardwood suitable for posts)

The modern Bikol equivalent of *alilintá'* is *lintá'*. There is no indication in the Lisboa dictionary, however, that *lintá'* is an independent form or an alternative to *alilintá'*.

Lisboa cites only one pair of words as related, the doublet in example (21) which shows what is apparently regressive assimilation of the second *t* of the root, assuming that *tagpó'* is indeed the word base for *alitatpó'*.

- (21) *alitatpó'/tagpó'* tree (type which produces a small, edible fruit)

Examples (22) and (23) show an obvious relationship between the affixed and unaffixed forms. In example (24) this relationship is less obvious.

- (22) *alinsunód* MAG-, -ON to follow behind or after  
*sunód* MAG-, -ON to follow behind; to occur after
- (23) *alintamóng* MANG- to wear a blanket or shawl across the shoulders  
*tamóng* blanket; MAG- to cover someone with a blanket
- (24) *alimpungay* MANG- to be at the uppermost part of a house or the highest reaches of a tree  
*pungay* MAG- to be filled to the point of overflowing

There is the possibility that *ali(N)-* is composed of a series of two affixes, *a-* and *li-*, the homorganic nasal later being added by epenthesis. In the discussion of the prefix *a-* (section 2), one of the doublets *alipúdok/lipúdok* indicated that this affix combination was a possibility.

4.2 THE PREFIX *li-*

The appearance of *li-* as a prefix for a few of the headwords in the Lisboa dictionary can be explained by reduplication of the first CV of the base.

- (25) *lilingóy/salingóy*      MAG-, -ON to turn the head to look at something behind  
       *lingóy-lingóy*            MAG-, -ON to turn the head repeatedly to look at something behind

There are many more examples where an apparent *li-* prefix is simply part of an affixed stem containing an *-in-* infix. Lisboa cross-references all of the examples presented below.

- (26) *linama'*                    fronds of the *huri* palm which are tinged with red  
       *lama'*                         MAG-, -ON to dye the fronds of the *huri* palm red
- (27) *linambán*                 room, cell  
       *lanób*                        wall; MAG-, -AN to enclose with a wall
- (28) *linanot*                    type of bread (made from wheat flour, resembling a tortilla, but thicker and harder)  
       *lanot*                         MAG-, -ON to make the type of bread called *linanot*
- (29) *linúgaw*                  type of rice porridge  
       *lúgaw*                        MAG-, -ON to make *linúgaw*

There are a number of entries, however, which cannot be so easily dismissed. Example (30) is identified by Lisboa as a doublet. The full entry for *lipu'ngáw* is interesting as well since Lisboa treats *li-* as a prefix which is deleted when *lipu'ngáw* is prefixed with *hiN-*. In other words *hiN-* + *lipu'ngáw* is not \**hilipu'ngáw* but *himu'ngáw* or *hipu'ngáw*.

- (30) *lipu'ngáw/pu'ngáw*      MA- to feel sad, lonely, melancholy

Example (31) is not cross-referenced by Lisboa, but a relationship between the pair is obvious.

- (31) *libutóg*                    MAG-, -ON to puff out the cheeks  
       *butóg*                        bladder; MAG- to fill with wind (the stomach); -AN to suffer with this affliction (a person)

The relationship between the pairs of words in examples (32) and (33) is not as obvious, yet it can be seen.

- (32) *ligpí'*                      the flat part of the forehead  
       *gapí'*                         MAG-, -ON to beat down grass or weeds to make an area suitable for hunting with nets
- (33) *lipárong*                 PANG-AN: *pangalipadngán* to have one's vision blurred (as from tiredness or glare)  
       *párong*                      MAG-, -ON to put out a candle or a fire; MA- to go out (a fire, candle); MA-AN: *mapadngán* to be left (as in the dark, in the cold) when a candle or fire goes out

Examining the entries in the Lisboa dictionary produced no possible derivational origins for the *li-* prefix. Example (33), however, indicates that this search was for the wrong affix. When *lipárong* is



prefixed with *paN-*, this prefix is not added to the word base (or stem, depending on the final analysis) *lipárong*, since the resultant affixed form would be *panlipadngán*. *PaN-* is prefixed to a base or stem of the form *alipárong*. The prefix *li-* can be explained as the resulting form after the loss of the initial vowel of a prefix of the form *ali(N)-*. The existence of the doublet *alipúdok/lipúdok* can also be used to strengthen this argument.

#### 4.3 THE PREFIX *ali(N)-* RECONSIDERED

There are only two related headword entries in the Lisboa dictionary to which a prefix such as *ali(N)-* is formally and semantically similar. These are *alin* and the reduplicated *alin-alin*. Both these entries deal with implied or overt comparison. The relevant parts of both of these entries are presented below.

- |                  |   |
|------------------|---|
| (34) <i>alin</i> | MAGPA-, (PAG-)PA-AN to compare items in a group for the purpose of deciding which one to choose; MAGPA-, I-(PAG-) to compare one item with others so as to decide which one to choose |
| <i>alin-alin</i> | a great variety (from which to choose); MAGPA- to be abundant in variety (items, from which a choice is to be made)   |

Given these entries, it is possible to interpret words affixed with *ali(N)-* as referring to something 'in comparison to' the base, for example *libutóg* meaning 'to puff out the cheeks' in comparison to the base *butóg* referring to the bloating of the stomach; *ligpi* 'the flat part of the head' in comparison to *gapí* 'referring to the flattening of an area of grass or reeds; and *lipárong* meaning 'blurred vision' in comparison to *párong* referring to extinguishing the light of a fire or candle.

*alin(N)-* is undoubtedly a very early prefix. Its widespread occurrence in Austronesian languages was referred to above. It also appears to have the freedom to combine with the affixes *ba-*, *sa-*, *ka-* and *ta-* (see sections 7, 11 and 13).

A combination such as *ma-* + *ali(N)-* resulting in a prefix of the form *mali-* might also explain Tagalog forms such as *malinamnam*, *maligamgam* and *maligasgas* meaning, respectively, 'tasty', 'anxiety' and 'the roughness of the skin'. It is described in one Tagalog grammar (Gana 1964:54) as having the same function as *mala-* meaning 'like' or 'similar to'. It is also described as a prefix old enough to have made reconstruction of the word base problematic. Only *malinamnam* is listed in Tagalog dictionaries under the entry *namnam*. *maligamgam* and *maligasgas* are listed under *ligamgam* and *ligasgas* respectively.

#### 4.4 THE PREFIX *ari(N)-*

There is a much smaller group of words which are prefixed with *ari(N)-*. Only one of these (35) can be shown to have a fairly provable correspondence with an unaffixed base using internal evidence. A number of other entries, *arituktók* 'the peak of a mountain' and *aripunsó* 'the crest of a hill or summit of a volcano' for example, look quite convincingly affixed if external evidence is used and a comparison is made to the Tagalog *tuktók* and *punsó* which have almost identical meanings.

- |                       |  |
|-----------------------|--|
| (35) <i>aripumpón</i> | a flock or herd; a group of people; MAG-, -ON to drive a flock; to herd together |
|-----------------------|--|

*pumpón* MAG-, -ON to gather what is drying in the sun; to collect fish with the hands; to gather together her chicks (a mother hen)

The prefixes *ari(N)-* and *ali(N)-*, in addition to their formal similarity, are actually given as alternative forms by Lisboa for one entry.

(36) *aringáhot/alingáhot* MA- to feel oppressively hot due to the clothes one is wearing  
*ari(N)-* is formally and semantically similar to the interrogative pronoun *arin* 'which', a word also indicating the same type of comparison discussed for *alin* above.

#### 5. THE PREFIXES *alu-* AND *aru-*

There are small groups of entries which appear to be prefixed, respectively, with *alu-* and *aru-*. Example (37) shows the one pair of entries where this correspondence can be seen.

(37) *alupating* MAKA-, MA- to overheard something, as when people talk in a whisper, or at a distance

*pating* partially deaf; MA- to be partially deaf

There are, unfortunately, no examples where a comparison between an entry affixed with *aru-* and an unaffixed form can be made. The case for an *aru-* affix, however, can be strengthened by comparing the following two affixed forms.

(38) *alusu'os* whirlwind; MAG- to form such a whirlwind  
*arusu'os* used to describe smoke which rises from a fire into the air

#### 6. THE PREFIX *ati-*

There are a number of headword entries in the Lisboa dictionary which appear to be prefixed with *ati-*. Lisboa, however, makes no mention of doublets and cross-references none of his entries.

Example (39) is a restatement of the entry for *atitiras* included in the discussion of the prefix *a-* (section 2). Example (40) shows the suffix *-(l)an* as well as the *ati-* prefix. The usual suffix for such words is *-han*.

(39) *atitiras/titiras* caterpillar (type which burns like fire when touched); MAG- to eat something (of such a caterpillar)

*tiras* MA- to burn or scald the mouth when eating something hot or chewing something piquant (such as an areca nut mixture)

(40) *atibungalan* the stem of the areca nut or coconut

*búnga* areca palm, producing a nut or fruit of the same name

Looking briefly at external evidence to strengthen the argument for the existence of such a prefix, the Bikol *atibángaw* 'botfly' is equivalent to the Tagalog *bángaw*. There is no independent root *bángaw* in Bikol. Tagalog does have an affix of the form *pati-* which is used to show sudden and unanticipated actions (Gana 1964:37, Blake 1925:258), for example, *magpatihód* 'to fall to one's knees' and *magpatihúlog* 'to suddenly and unexpectedly fall'. For the purposes of this paper, such

an affix would have to be analysed as comprising two parts: *pa-* + *ati-*. *pa-* could easily be interpreted in its normal function as causative. That still leaves the question as to the meaning of *ati-*.

Returning to Bikol, neither the origin nor the meaning of this affix is particularly clear. The remaining entries where *ati-* is clearly indicated as a prefix are presented below.

(41) <i>atibuag</i>	MAG-, I- to throw something which one has on the palm of one's hand, on a plate, etc.
<i>atiku'al</i>	MAG-, I- to throw something with the hand, in anger or annoyance
<i>atipalo</i>	maggots which breed in cooking stoves
<i>atipolo</i>	type of tree (large, producing a fruit with edible <i>picos</i> )

## 7. THE PREFIXES *ba-* AND *baN-*, AND THE INFIX *-al-*

### 7.1 THE PREFIXES *ba-* AND *baN-*

There is also evidence for a prefix of the form *ba-*. In the first set of examples below, *ba-* most likely arises from reduplication. These entries, which also contain the locative suffix *-an*, are all cross-referenced by Lisboa.

(42) <i>babaháwan</i>	a place for storing leftover food
<i>báhaw</i>	leftover food
(43) <i>babayongan</i>	pigsty; enclosure for keeping animals
<i>bayong</i>	an animal, such as a pig, kept in an enclosure for the purpose of being fattened
(44) <i>babató'an</i>	a small box or pouch with a flap used for storing weights called <i>bató'</i> used for measuring gold
<i>bató'</i>	weights for measuring gold
(45) <i>baba'singan</i>	weights of one tael, used for measuring gold
<i>ba'sing</i>	the weight of one tael

Examples (46) and (47) are not cross-referenced by Lisboa, although the same type of locative relationship discussed above is apparent.

(46) <i>babaghán</i>	waist
<i>bagá</i>	spleen
(47) <i>babalónan</i>	gullet, gizzard of fowl
<i>bálon</i>	food which is carried from the house, later to be eaten on a trip, in the fields, etc.

Examples (48) and (49) are non-locative and contain no *-an* suffix. It is unclear, however, whether *ba-* in these instances arises simply from reduplication of the first CV or is independently affixed to the root following the pattern discussed below.

- (48) *babatid* MAG-, -ON to sound someone out; to try to find out what has happened  
*batid* MAG-, -ON to test something; to try something out
- (49) *babasá'* MAG-, -ON to wet, moisten  
*basá'* MAG-, -ON to wash off; MAG-, -AN to wash something

Lisboa does cite one doublet:

- (50) *baku'ong/ku'ong* MAG- to struggle under a heavy weight carried on one's shoulders or back

There are other entries, however, where a relationship between a *ba-* affixed and unaffixed form seems clear.

- (51) *baklás* MAG-, -ON to remove bark from a tree, skin from the body  
*kalas* MAG-, -ON to remove meat from the bone
- (52) *bakurong* a rope snare used for catching the birds called *máya* and *dignós*  
*kurong* a bird in a cage; MAG-, -ON to keep a bird in a cage
- (53) *balnáng* MAG- to glisten or shine (as gold, silver, the water of the sea)  
*linang-linang* MA- glistening, shining (a smooth or polished surface)
- (54) *balitok* referring to something twisted, like the horn of a ram; MAG-, -ON to twist something in this way  
*litok* MAG-, -ON to go to get something in a roundabout way, and not by going straight to a destination

Although *ba-* in example (55) appears as a probable prefix using Bikol evidence alone, *balising* may also be related to the Malay *bising* 'noisy' with an infix *-al-* (see section 7.2).

- (55) *balising* MAGKA-, PAG-AN to misunderstand what one hears; to not know what to make of what one hears  
*lising* MAG-, -ON to disturb someone by making a noise

In example (56) the root is both prefixed with *ba-* and suffixed with *-an*.

- (56) *bagtawan* type of fish (flat, colourful, resembling a disk)  
*gataw* MAG- to float, something in the middle of the water, such as a crocodile or a fish showing its head or another part of its body above water

Examples (57)-(59) form an interesting subgroup since both the affixed and unaffixed forms are nouns.

- (57) *balá'og* gift; the desire to see someone (the modern meaning is 'sincere' or 'heartfelt')  
*la'óg* the inside
- (58) *balf'og* necklace, collar

	<i>l'og</i>	neck
(59)	<i>bantu'lang</i>	bracelet (type worn by men)
	<i>tu'láng</i>	bone

The occurrence of the homorganic nasal in example (59) raises the question of whether there are two affixes, one of the form *ba-* and one of the form *baN-*. Having two affixes of similar form would be parallel with affix sets such as *sa-*, *saN-* and *ta-*, *taN-* discussed later in this paper. In the discussion of the prefixes *sa-* and *saN-* (section 11), both forms can be shown to come from distinct sources. With *ta-* and *taN-* (section 13) the situation is less clear. Since *ba-* and *baN-* both appear to indicate similarity, and *baN-* is fairly restricted in its occurrence, it may be best to consider these a single affix of the form *ba-* with the homorganic nasal added through epenthesis.

One further example with *ban-* follows. Example (60) presents a series of entries which show a derivational chain.

(60)	<i>bantulin</i>	MAG-, I- to tell another something that has been said to you
	<i>tulin-tulin</i>	MAKA- to be the talk of the town; to spread from person to person (news)
	<i>tulin</i>	MA- swift (a boat); MAG- to travel swiftly

The most consistent meaning attributable to *ba-* is 'likeness' or 'similarity'. This may be related to the Bikol *bagá* which has a similar meaning, for example *ba + basá*, 'like washing', *ba + kalas* 'like removing meat from the bone', *ba + kurong* 'like keeping a bird in a cage', *ba + lináng* 'like a glistening polished surface' etc.

A number of entries which appear to be prefixed with *ba-*, however, cannot be explained by the solutions offered above. These entries will be looked at below.

## 7.2 THE INFIX *-al-*

A number of the entries examined to determine if they contained the prefix *ba-* revealed another pattern, that of the infix *-al-*. A cognate affix exists in other languages of the Austronesian language family – in Malay, for example, where, even though it is fossilised, the infix *-el-* (where *e* represents the *pepet* or *schwa*) still manages to rate a mention in most grammars (Asmah 1986:61). Compare Malay *sidik* and *selidik* meaning 'to examine in detail' and 'to investigate' respectively. The following examples are from the Lisboa dictionary.

(61)	<i>balakbák</i>	the bark of plants such as the banana or abaca in which the trunk itself is made of such bark-like layers
	<i>bakbák</i>	MAG-, -ON to peel the bark from a tree or from plants with a similar covering
(62)	<i>balantók</i>	MAG-, -ON to throw a knife, or dagger, or similar object, holding it by the point
	<i>bantók</i>	MAG-, -ON to strike or wound someone with the back edge of a knife blade (not the side that is sharpened)

- (63) *balagat* MAG-, -ON to run out after something, so as to head it off and capture it as it runs toward you  
*bagat* MAG-, -ON to go out to meet those passing by on the road or river, or to call or speak to them
- (64) *balutbót* MAG-, -ON to remove the cargo from a ship in order to reload it more satisfactorily or to see if it has become wet  
*butbót* MAG-, -ON to remove something that has been hidden, put away for safe keeping, or placed between other items
- It is not, however, only entries beginning with *b* that show such an infix. The following examples were discovered while investigating possible prefixes of the form *ta-* and *sa-*.
- (65) *salakay* MAG-, -AN to ship out on a particular vessel  
*sakáy* MAG-, -AN to board a ship
- (66) *salakat* MAG- to come out from the mud of the rice fields, moving toward the irrigation canals or the higher reaches from where water enters the fields in order to spawn (of fish following heavy rains marking the end of the dry season)  
*sakát* MAG- to climb up
- (67) *salaghóm* MAG- to be absorbed (liquid)  
*saghóm* MAG- to be absorbed (liquid)
- (68) *salúno'* MAG- to follow one after the other to a particular destination (two people)  
*súno'* MAG-, -ON to follow in the footsteps of one who has gone before; to repeat a refrain that one has said before (as when rowing)
- (69) *talapi'* boards which are placed on the sides of boats; MAG-, -ON to fix such boards to the sides of boats  
*tapi'* board, plank; MAG-, -ON to manufacture or produce a board
- (70) *taluntón* rattan or cord which is lowered, enabling people to then climb up; MAG-, I- to lower such rattan or rope for this purpose  
*tuntón* MAG-, I- to lower something suspended in air, such as a lamp at the end of a rope

It is not easy to identify a unified meaning for an affix such as *-al-* from the data presented. In Sundanese and Javanese an infix of form *-Vl-* or *-Vr-* generally carries the meaning of recurrence or plurality, and it is possible to interpret such meaning in the fossilised Malay examples as well (Mees 1969:203-205). Plurality or recurrence in Bikol is shown by an infix of the form *-Vr-*. The active Bikol plural infix *-Vr-* and the fossilised infix *-Vl-* are probably related. The Bikol prefix *pára-* and the Tagalog *palá*, both showing habitual or repetitive action, must also be related. The affix here is probably *pa-* showing, respectively, *-ar-* and *-al-* infixes. What is not possible on the basis of the



Bikol data presented is to determine an acceptable meaning for the infix *-al-*, and so it is best to move on to other considerations.

#### 8. THE PREFIXES *ba-* + *ali(N)-* AND *balik-*

A glance at the entries in the dictionary beginning with *bali(N)-* and *balí(k)* indicates that an affix or affixes of this type once existed. Many of these entries, either literally or figuratively, share the meaning of being 'turned around' or 'turned upside-down'.

- |                         |   |
|-------------------------|---|
| (71) <i>balikwá'</i>    | MAG-, -ON to turn something face up   |
| (72) <i>balikwág</i>    | MAG-, -ON to place what is on top on the bottom, and what is on the bottom, on top (as when one mixes food or rice in pot); variant <i>baliswág</i>                   |
| (73) <i>balikwít</i>    | MAG-, -ON to lift someone from under the arms   |
| (74) <i>baligbíg</i>    | MAKA- (MA-) to catch something about to fall; MAKAMANG-, MAMANG- to repair something that is broken; to turn someone around (as from behaving badly to behaving well) |
| (75) <i>baligwát</i>    | MAG-, -ON to lever something into a position where it is at an equal height with something else; variant <i>bali'at</i>   |
| (76) <i>balingag</i>    | MAG-, -ON to overwhelm; to take someone by surprise   |
| (77) <i>balintóng</i>   | MAG-, -ON to turn what is face up, face down, and what is face down, face up  |
| (78) <i>baliskád</i>    | MAG-, -ON to turn something right side up; to roll something over on its other side   |
| (79) <i>balisungóng</i> | funnel-shaped; MAG-, -ON to form something into such a shape  |
| (80) <i>balisot</i>     | upside-down, inside-out; MAG-, -ON to stand on one's head; to turn something inside-out   |

If there are two different affixes here, the first, *bali(N)-*, is probably a combination of two prefixes previously identified, that is *ba-* + *ali(N)-*. This argument is strengthened by the existence of the affix combinations *sa-* + *ali(N)-* and *ta-* + *ali(N)-* (see sections 11 and 13). The cumulative meaning of this affix sequence would be 'in comparison to' or 'like'.

The second, *bali(k)*, is probably the base *balik* which during Lisboa's time had the basic meaning of 'reversed' or 'inside-out' and the figurative meaning of 'returning to a place where one has been'. A more decisive conclusion, however, regarding these affixes cannot be reached on the basis of internal evidence found in the Lisboa dictionary.

#### 9. THE PREFIX *hiN-*

Although the prefix *hiN-* in modern Bikol is fossilised, there is evidence in the Lisboa dictionary to show that *hiN-* was far more widely used in the past. *hiN-* appears to mark the transition from one state to another. This interpretation alone is able to account for its rather inconsistent marking of both the inception and conclusion of actions, and the loss or restoration of particular states.

- (81) *bálon* stores or provisions of food or money which are carried on a trip  
MAG-, -ON to take as stores or provisions; MAHING-:  
*mahimálon* to be used up (stores or provisions); MAHING-AN:  
*mahimalonan* to be left without stores or provisions;  
MAKAHING-: *makahimalon* to cause the loss or lack of stores  
or provisions
- (82) *puró* point or tip; MAG-, -AN to put a point or tip on; MAG-HING-,  
HING-ON: *maghimuró*, *himuró* to cut off the tip MAG-HING-,  
HING-AN: *mahimuró*, *himuróhan* to cut the tip or point from
- (83) *mutá'* the mucus that collects in the corner of the eyes; MAG- to have  
one's eyes noticeably filled with such mucus; MAGHING-,  
HING-ON: *maghimutá'*, *himuta'ón* to remove such mucus
- (84) *pulá* the colour red; MA- red; MAG- to turn red; MAHING-:  
*mahimulá* reddish; showing signs of red; MAGHING- to begin  
to turn red
- (85) *ráhay* MA- good; MAG- to improve; to get better; MAG-, -ON to put  
something away for safekeeping; to keep safe, secure;  
MAGHING-, HING-ON: *mighiráhay*, *hiraháyon* to repair; to  
restore to good condition
- (86) *tagbó'* MAG-, -ON to meet someone; MAGHING-, HING-AN:  
*maghinagbó'*, *hinagbo'án* to wait for the opportunity to meet  
someone

#### 10. THE PREFIXES *mu-* AND *pu-*, AND THE INFIX *-um-*

During the investigation into *hiN-* as an affix, a number of forms were discovered which were not immediately explainable in terms of *hiN-* being prefixed directly to an identifiable word base. *hiN-* instead appeared to be affixed to a stem, or to a base which had been previously derived or inflected. Example (87) is one such entry.

- (87) *himurangos/himurangsá'* referring to words said in anger, such as when a donkey kicks its  
packsaddle; MANG-, PANG-ON to blame someone who is  
without fault

Example (87) has the following modern equivalent:

*murangsá'* MAGPA-, PA-AN to take anger out on one who is without blame

While the entries above are obviously related, *murangsá'* is not given as an entry in the Lisboa dictionary. Using diachronic comparison of the two entries, it is obvious that the prefix here is *hiN-*. *murangsá'*, being trisyllabic, is not a common root form, but no disyllabic root was discovered. An additional problem is accounting for what appears to be a prefix of the form *mu-* or *pu-*, as the sequence *hiN-* + *purangsá'* would result in a word of the form *himurangsá'*. There were other entries as well which showed such a prefix.

- (88) *himurayá'* MAG-, -AN to learn something from experience

- rayá'* MAG-, -ON to teach someone a lesson; to punish someone to prevent him from repeating a particular act
- (89) *himurá'ot* MAG-, PAG-ON to keep watch over someone who is near death
- rá'ot* MA- bad; MAG- to go bad; to spoil

There is still a question as to the form of the prefix, whether it is *mu-* or *pu-*, or, indeed, if there are two prefixes which must be explained.

### 10.1 THE PREFIX *pu-*

Lisboa gives cross-references for three sets of entries, presented as examples (90)-(92). He, however, does not make reference to a possible base. For example (92) no possible word base was found. The form in common use in modern Bikol is *kurá'mos*. The status of *ku-* as a formative in such constructions has not been investigated.

- (90) *mulingáw/pulingáw* MANG-, PANG-ON to request the settlement of debts from the heirs of an estate
- lingáw* MA- to forget; MA-, MA-AN to forget about something
- (91) *murapok/purapok* MAKA-, MA-AN to burden someone with the responsibility of doing everything
- rapok* MAG-, I(PAG)- to cook a lot of taro on the coals of a fire; MAG-, (PAG-)-AN to cover the coals of a fire with a lot of taro; (metaphorically) PAG-AN to be told a pack of lies: *Pinagrapokan akó kainíng anán mga mará'ot* 'I am dumped on by everything bad' (Lisboa gives this the figurative interpretation of 'What I am told is worthless')
- (92) *mura'mos/purá'mos* MAG- or MANG- to wash one's face; MAG-, -ON or MANG-, PANG-ON to wash dirt from the face

It is uneconomical to consider two separate affixes when there are active processes which might have resulted in one form having being derived from the other. *maN-* + *pulingáw* results in *mamulingáw* and *maN-* + *purá'mos* results in *mamurá'mos*. If the meaning attributed to the independent affix *pu-* was gradually lost it would have been considered a part of the root. *ma-* would have been recognisable as an independent form by analogy with the previously existing prefix *ma-*, and the remaining forms would have gradually come to be considered the roots, resulting in *mulingáw* and *murá'mos* respectively.

There is some additional evidence to show that *pu-* was an independent affix, although its origin and meaning are not possible to ascertain using internal Bikol evidence alone. Examples (94) and (95) are also suffixed with the locative *-an*. In example (95) it is not possible to determine whether *pu-* is an independent affix or the result of reduplication of the first CV of the base.

- (93) *pugapó'* type of bird (mottled, usually found walking among stones)
- gapó'* stone
- (94) *pulayágan* mast; MAG- to possess a mast (a boat); MAG-, -AN to fit a boat with a mast

- láyag* sail; MAG- to navigate a boat with a sail  
 (95) *pupulóan* handle of a knife  
*púlo* handle of a bolo knife; MAG-, -ON to fit a handle; MAG-, -AN to fit with a handle

Example (96) is presented as a doublet by Lisboa. The form *pusngák* most probably results from the processes of assimilation and deletion applied to the sequence *pu- + usngák*.

- (96) *pusngák/usngák* MAG- to expel air quickly through the nostrils (as one might when a mosquito flies in or when water enters when swimming)

It is also interesting to note that the Bikol *purísaw* referring to a 'troubled or disturbed sleep' appears in Malay as *risau* meaning 'anxiety' or 'apprehension'.

## 10.2 THE PREFIX *mu-* AND THE INFIX *-um-*

The prefix *mu-* is easier to explain in terms of other affixation than *pu-*. Examples (97) and (98) can be explained in terms of back formation from roots affixed with *maN-*. *mu'ót* 'to love' and *mútik* 'to lie' would then originate, respectively, from the roots *bu'ót* and *pútik* in the following way: *maN-* + *bu'ót* > *mamu'ót*, *mu'ót* and *maN-* + *pútik* > *mamútik* or *mútik*.

- (97) *mu'ót/bu'ót* desire, will; MANG-: *mamu'ót* to like, love  
 (98) *mútik/pútik* a lie; MANG-: *mamútik* to lie

Other entries in which one member of a doublet cited by Lisboa begins with an initial *mu-* may also be explained in terms of back formation. The original affixation for examples (99) and (100), however, would be different from that postulated for (97) and (98).

- (99) *munínong/tunínong* MA- capable, discreet, cautious; MAG- to become discreet, cautious  
 (100) *musang/usang* MAG- to have the mouth filled with food

An infix of the form *-um-*, now an alternative command affix, functioned differently when Lisboa was compiling his dictionary. Even at this time the affix must have been in marginal use since it is rarely cited as regular affixation for his entries. In examples (101) and (102) the full form of the verb is given in parenthesis and the *-um-* infix is bolded.

- (101) *Anay dumtóng (dumatóng) akó, nagháli' siyá.*  
When I arrived, she left.  
 (102) *Anay aso kumháng (kumaháng) akó...*  
When I went...  
 (103) *saró'* one; MAG-, -AN or MAKI-UM-, PAKI-UM-AN: *makisumaró'*, *pakisumaró'án* to join together with  
 (104) *sáyod* MAG-, -ON to explain or clarify; MAKI-UM-, IPAKI-UM-: *makisumáyod*, *ipakisumáyod* to return to ask what was previously mentioned so that it can later be accurately reported



- (109) *sahuri* MAG-, -ON to punish or reprimand someone after he has stopped interfering or being a nuisance  
*huri* last; following behind; MAG- to stay behind
- (110) *salingóy* MAG-, -AN to look behind at something by turning the head slightly  
*lingóy-lingóy* MAG-, -ON to constantly turn the head to look at something behind
- (111) *sawaló* rope made from eight strands; MAG-, -ON to make such rope eight  
*waló*

Regarding example (112), the modern meaning for *la'ág* is 'to put' or 'to place'. Lisboa does not cite this as one of his entries.

- (112) *sala'ág* MAG- to not have a definite nest, and so lay eggs in different places from day to day (a hen)  
*la'ág-la'ág* MAG- to go from town to town without residing in any particular place

## 11.2 THE PREFIXES *sa-* + *ali(N)-*

There are a number of entries which can most easily be analysed as a sequence of the two prefixes *sa-* + *ali(N)-* (see section 4), the prefix *sa-* setting the locative or possessive relationship, and the prefix *ali(N)-* the comparative one. The clearest indication that *sa-* and *ali(N)-* are independent elements can be seen by the variations possible for example (113).

In addition to the entries *salipapaw* and *salimpapaw* there is a further entry *kalimpapaw*. Assuming the word base is *papaw*, the affixed forms can be shown to be built up from the affix and base sequences: *sa-* + *ali(N)-* + *papaw* and *ka-* + *ali(N)-* + *papaw*. In reality, no base of the form *papaw* is listed in the Lisboa dictionary. The closest base is *pawpaw*, and this is the base form shown in example (113). Lisboa, however, does mention the alternative base form *kapaw*.

- (113) *salipapaw/salimpapaw* the upper part; the surface area; MAG-, -ON to remove the upper part; to remove an outside layer  
*pawpaw* MAG-, -ON to remove an excess amount from the top; MAG-, -AN to skim the top of an excess amount

Similar relationships of association and comparison can be seen in the following examples.

- (114) *salibábaw* referring to something which is placed above something else; MAG-, I- to place something above; MAG-, -AN to cover something with something placed above it  
*bábaw* I- on, over, above; MAG+I-: *magibábaw* to climb up; to go up; MAG+I-, I+I-: *magibábaw, iibábaw* to place something on top or over
- (115) *salibong* MAG-, -ON to run from one place to another, as when chasing after something in an attempt to catch it



<i>ibóng</i>	the far side of a river, road; the other side; MAG-, -ON to do something with both feet or hands; to work on both sides of something
(116) <i>salimagyó</i>	a large storm, not as strong as a <i>bagyó</i> ; MAG- to strike or develop (such a storm); -AN to be affected by such a storm (a person, place)
<i>bagyó</i>	typhoon; MAG- to develop (of a typhoon); -ON to be destroyed by a typhoon; -AN to be affected by a typhoon (a person, place)

11.3 THE PREFIX *saN-*

There are two prefixes of the form *sang-* which appear as individual headword entries in the Lisboa dictionary. The first of these is a combining form meaning 'one'. Lisboa relates this form to the common word for 'one' in Bikol, *saró*, although it is more probably derived from *isá* which was restricted during the time of Lisboa to counting: *isá*, *duwá*, *tuló* etc. 'one, two, three' etc. Although the combining form *sang-* continues to be used, *isá* has completely disappeared from modern Bikol. *sang-* is clearly derived from a sequence of *isá* and the linker *-ng*.

All of the numbers in example (117) can be accounted for in this way. Lisboa gives an alternative form for *sanglaksá'* which shows assimilation of the nasal, *sanlaksá'*. Modern Bikol only uses the assimilated form of *sang-púlo'*, which is *sampúlo'*.

(117) <i>sangpúlo'</i>	ten
<i>sanggatós</i>	one hundred
<i>sangribo</i>	one thousand
<i>sanglaksá'</i>	one million

The second of these prefixes means 'as many as' or 'as numerous as', and is seen by Lisboa to be related to the other *sang-* prefix discussed above. Some of the examples presented by Lisboa are shown as example (118).

(118) <i>sangdagá'</i>	as endless as the land
<i>sangbaybáy</i>	as numerous as the grains of sand
<i>sangdáhon</i>	as plentiful as the leaves of a tree

In examples (119) and (120) the prefix *saN-* indicates 'oneness' or 'unity'. Example (119), for instance, can be interpreted as 'of one blood', and example (120) as the union of two people working together or focusing on the same chores. The modern form of *sanggáwad* is simply *gáwad*.

(119) <i>sangdugó'</i>	blood brothers; the relationship formed when blood is taken from two people, mixed with local wine or <i>tubá'</i> and then drunk by both from the same vessel
<i>dugó'</i>	blood of an animal; MAG- to bleed
(120) <i>sanggáwad</i>	MAG-, -AN to help someone with ordinary work in the house, or similar everyday chores, but not with work in the fields; MAG- to help one another in such tasks; MAG-, PAG-AN: <i>magsanggáwad</i> ; <i>pagsanggawáran</i> to work on a particular task together; KA- helper, assistant

*gáwad* MAG-, -ON to take something down from a place higher than where one is positioned; to unload cargo; MAG-, -AN: *maggáwad, gawáran* to take something down from; to unload cargo from

In example (121) it is not clear if a boatman becomes 'one' with the shore, favouring an interpretation of *saN-* for the prefix, or if the shore is simply to be interpreted as the location of the action, favouring *sa-* as the prefix.

(121) *sangílid* MANG- to navigate a boat along the bank of a river and not in the centre

*gilid* margin or edge; the bank of a river or shoreline of the sea; *sa gilid nin dágat* the shore of the sea; MAGPA-, PA-AN to head toward the bank or the shore

The prefix *saN-* in example (122) most likely takes on the secondary meaning of 'as numerous as', with the closing of the eyelids of one having trouble staying awake, compared to the flickering of a candle about to go out. The modern meaning of *piráw-pitáw* is 'to blink the eyes (as if trying to remove something that has fallen into them)'.

(122) *sampiraw* MAG- to be unable to keep one's eyes open (a person falling asleep)

*piráw-pitáw* MAG- to flicker (as a torch or candle about to go out)

In example (123) the meaning of *saN-* is broadened from 'one' to that of 'same'. The stem identified is itself morphologically complex, being built on the base *búlaw*, which refers to a particular measure of gold.

(123) *sambuláwan* type of beetle (yellow in colour)

*buláwan* gold

## 12. THE PREFIXES *taga-* AND *tagu-*

Lisboa lists both *taga-* and *tagu-* as headword entries. Although both of these prefixes are recognisable in modern Bikol, only the first retains some potential for use. This may be due to the influence of Tagalog where it is commonly used as a formative for agentives, such as in *tagálúto* 'cook'. Although *tagálúto* is perfectly acceptable in modern Bikol, the more common agentive formative is *para-*: *parálúto*'.

The prefix *taga-* can be seen as related to the word *tagá*, which associates a person with the place he or she is from – *Tagá Cana'mán akó* 'I'm from Cana'mán' – and *tagu-* to the word *tágo*, which means 'to hide'. As a prefix *tagu-* deals with intrinsic qualities or attributes.

### 12.1 THE PREFIX *taga-*

Lisboa gives three meanings of *taga-*. The first, as a prefix to verbs, indicates the frequent, though not consistent, performance of the action of the root. This is related to the Tagalog-influenced function discussed in the introduction above. The second meaning (related to the first), as a prefix to nouns, indicates the common possession of, or association with, the root indicated. The third,

generally affixed to nouns representing parts of the body, indicates the level to which something has reached.

- (124) *bása* MAG+TAGA-: *magtagabása* to read frequently  
 (125) *tumbák* MAG+TAGA-, PAG+TAGA-ON: *magtagatumbák*,  
*pagtagatumbakón* to walk around armed with a lance  
 (126) *tuhód* TAGA-: *tagatuhód* reaching up to the knee

Examples (127)-(129) show *taga-* as their prefix. They are not cross-referenced by Lisboa, although the entry shown as example (129) has a recognisable root in *súso* 'breast'. The additional prefix *ma-* is probably adjectival. *taga-* in example (129) is most closely related to the first meaning discussed above. Examples (127) and (128) have derived either from the first or second meanings indicated.

- (127) *tagaltáw* buoy; MAG-, -AN to mark with a buoy  
*latáw* MAG- to float  
 (128) *tagalpó'* a herb used by hunters to charm game so they will not flee;  
 MAG-, -ON to charm game in such a way; MAG-, I- to use a  
 herb for such a purpose  
*lapó'* dislocated, broken (a limb); MAG-, -ON to break or dislocate the  
 limb of an animal  
 (129) *tagamasúso'* midwife; MAG-, PAG-AN to aid a woman in giving birth (a  
 midwife)  
*súso'* breast, teat; MAG-, -ON to suck milk from the breast (a child);  
 MAGPA-, PA-ON to nurse a child (the mother)

## 12.2 THE PREFIX *tagu-*

Lisboa defines *tagu-* as a prefix which is added to abstract roots to indicate the qualities by which something is considered or appraised.

- (130) *ráhay* good; MAKA+TAGU-: *makataguráhay* to be considered as  
 reasonable and good  
 (131) *rá'ot* bad; MAKA+TAGU-: *makatagurá'ot* to be appraised as being bad  
 or unsuccessful

This general meaning can be seen in the entry pairs (132)-(134) below which are not cross-referenced by Lisboa.

- (132) *tagulá'in* conjecture or suspicion; a gut feeling; MAG-, -ON to have a gut  
 feeling about something; to suspect  
*lá'in* different from the others; MAG-, -ON to separate things out  
 according to differences; KA- strange, different: *Kalá'in mo doy*,  
 How strange you're acting.

- (133) *tagulipód* a root said to be able to render things invisible; MAG-, -ON to make something invisible using such a root; -AN: *si tagulipódán* that which is rendered invisible
- lipód* MAG-, -AN to block something from view by placing something in front of it
- (134) *tagulnód* a weight attached to something to make it sink; MAG-, -AN to place such a weight on something
- lúnod* MAG- to sink; MAGPA-, PA-ON to sink something

### 13. THE PREFIXES *ta-* AND *taN-*

The prefixes *ta-* and *taN-* are treated separately here mainly because they show different types of assimilation; *ta-* shows no assimilation to the initial consonant of the root, while *taN-* shows the nasal assimilation expected. The same type of situation was found with the affixes *sa-* and *saN-* (see section 11) and *ba-* and *baN-* (see section 7). The decision to treat these affixes as separate is different from the decision taken on *ba-* and *baN-*, where it was possible to show the affixes as having a similar meaning. With *ta-* and *taN-* this is not possible.

#### 13.1 THE PREFIX *ta-*

Three of the examples where *ta-* is the prefixed element relate to shape or nonlinear direction. *ta-* in these examples appears to function like the prefix *ha-*, which derives adjectives showing height, length and depth.

- (135) *taglíd* describing something turned on its side; MAG-, to be turned or lying on its side; MAGPA-, PA-ON to turn something on its side
- gilíd* describing something inclined or twisted to one side; MAG- to incline toward one side; to lean or be bent toward one side; MAGPA-, PA-ON to incline, bend or twist something to one side
- (136) *talikód* describing someone with his back turned; MAG- to turn one's back; MAG-, PAG-AN to turn one's back on
- likód* the back
- (137) *tablóg* describing something round; MA- round; MAG- to become round; MAG-, -ON to make something round
- bilóg* entire, complete; MAG-, -ON to complete something in its entirety

Examples (138) and (139) cannot be interpreted in the same way as examples (135)-(137). Here the affixed and unaffixed forms show little or no difference in meaning or function.

- (138) *tabgáw* MAG-, -ON to frighten or chase away an animal
- búgaw* MAG-, -ON to frighten or drive off birds or animals

- (139) *takulod* highlands planted to sweet potatoes, palms and other plants  
*kulod* hills or highlands used for cultivation

There is one example in which the sequence of prefixes *ta-* + *ali(N)-* (see section 4) is affixed to a root.

- (140) *talinduwá* MAG-, -ON to take two parts (of something divided into thirds)  
*duwá* two

### 13.2 THE PREFIX *taN-*

While the existence of the affix *taN-* is obvious as one reads through the Lisboa dictionary, the meaning of such an affix is hard to determine. Lisboa makes only one reference to a *taN-* affixed and unaffixed form, and this is presented as example (141). It is unfortunate that the dictionary shows no independent entry under the unaffixed form.

- (141) *tangpos/apos* MA-: *matangpos*; *Natatangpos nin gumagá'od ining sakáyan*  
 This boat is filled with rowers (almost to the point of sinking)

The following examples are presented to show further possible relationships between a *taN-* affixed and unaffixed root.

- (142) *tambubong* referring to one who lives in another village, eating and living alone for the whole time he remains away from his home village;  
 MAG- to live apart in this way  
*bubong* straw used to cover the roofs of houses; MAG-, -ON to cover the roof of a house with such straw
- (143) *tangkaway* the shoot or runner of a vine or similar plant  
*kaway* MAG- to put out roots that are too high to be useful in transplanting and therefore unable to be used (rice in a seedbed flooded with too much water)

### 14. THE PREFIXES *so-* AND *no-*

The prefixes *so-* and *no-* refer to time – respectively past and future. The affixes are not currently productive nor were they productive when Lisboa was compiling his dictionary.

The prefix *so-* comes quite clearly from *aso* which is used to show past time, as in *Aso naggúhit akó...* ‘When I wrote...’ or ‘When I was writing...’. The modern form is *kasó*, which probably derives from a sequence of *kan + aso*.

The origin of *no-* is not so clear. It most likely comes from *anó*, which basically means ‘what’. Lisboa cites *anó* as having the additional meaning of questioning ‘at what stage’ or ‘to what point’ something had reached in its progress or development. It does not have that meaning now.

- (144) *Anó na an simbáhan nindó?*  
 At what stage is (the building of) your church?

(145) *Anó na an pároy nindó?*  
How far along is your rice crop?

(146) *Anó na an aldáw?*  
What time is it? (lit. At what stage in the day is it?)

Examples (147)-(155) show parallel sets of word bases affixed respectively with *so-* and *no-*, followed by the relevant word base. For example (147), no word base of the form *anoy* could be found. The closest base is *ánay*, for which *áney* is given as an alternative form. The occurrence of the diphthong *ey* in *áney* is found nowhere else in Bikol, and is probably an error. It is only possible to speculate that this should, indeed, be *ánoy*.

(147) *soánoy* a long time ago  
*noánoy* in the future  
*áney* when

(148) *soarín* when (past)  
*noarín* when (future)  
*arín* which

In example (149), the notional word base *udmá'*, which can be isolated using the criterion of recurrence or contrast, has no independent meaning in either old or modern Bikol.

(149) *soudmá'* yesterday  
*noudmá'* tomorrow  
*\*udmá'*

(150) *sosaró'* the day before yesterday  
*nosaró'* the day after tomorrow  
*saró'* one

The definition Lisboa gives for *sorayó'* is somewhat suspicious, particularly when compared with the meanings of *sosaró'* and *norayó'*. It is possible to speculate that the meaning intended is 'any number of days before yesterday'.

(151) *sorayó'* the day before yesterday  
*norayó'* four days from now  
*rayó'* HA- far away

The forms *soróng* and *noróng* in example (152) can only be explained as originating from sequences of morphemes such as *so-* + *saró* + *-ng*.

(152) *soróng ta'ón* last year  
*noróng ta'ón* next year  
*saróng ta'ón* one year

There are no further examples with *no-*. The remaining examples prefixed with *so-* conclude this section.

(153) *soba'gó* earlier, recently  
*ba'gó* new

(154) *sohápon* yesterday afternoon  
*hápon* afternoon



- (155) *sobanggi* last night  
*banggi* night

15. THE INFIX *-imin-*

Although the infix *-imin-* is frequently cited by Lisboa in his dictionary, it has completely disappeared from modern Bikol which, however, has the infix *-umin-* which is not cited by Lisboa. The prefix *-imin-* functions to show completed states of action, being or existence.

- (156) *arahit* perfect, complete; extensive in application or knowledge; -IMIN-: *Iminarahit sindáng mangalagkalág*. They looked everywhere.
- (157) *ági* MAG- to pass through/by; -IMIN-: *Iminági lámang sa girumdóm ko*. It just happened to pass through my mind.
- (158) *ku;á* -IMIN-: *Atí ta iká an kiminu'á*. There is a suspicion that you are the one who took it.
- (159) *gibo* -IMIN-: *Subot si kuyán an giminibo kaiyán*. They say that it was so-and-so who did that.
- (160) *arat* the smell of smoke; -IMIN-: *iminarat* to smell of smoke
- (161) *sálog* -IMIN-: *siminálog* to live by the river

## 16. CONCLUSION

This paper examined the data available in the Marcos de Lisboa *Vocabulario de la lengua Bicol* in an attempt to discover various patterns of morphemes that might have previously served as affixes in Bikol. Identification was based on synchronic evidence internal to Bikol, although diachronic evidence was cited, and reference was made to languages other than Bikol to strengthen the argument being made.

I found strong supporting evidence for the existence of a prefix of the form *ali(N)-* which has generally been assumed to exist for Austronesian languages. I also found that this affix could occur in combination with other affixes, namely *ba-*, *sa-*, *ka-* and *ta-*.

Evidence was also examined for the existence of other affixes. The prefixes *alu-/aru-*, *ati-* and *pu-* can be proved to have existed on the basis of recurrence and contrast, although it is not possible to assign consistent meanings or to locate possible origins for these forms.

Parallel nasal-final and non-nasal-final affixes of the forms *ba-/baN-*, *sa-/saN-* and *ta-/taN-* were examined. No consistent decision has been reached on the treatment of such pairs. It is obvious that *sa-/saN-* were separate and traceable to distinct origins. It is not so obvious whether *ba-/baN-* and *ta-/taN-* were as distinct.

The prefixes *no-* and *so-* serve the same function in modern Bikol as they did in late sixteenth and early seventeenth century Bikol – that is, to indicate future and past time respectively. Examining the entries in the Lisboa dictionary, however, made it possible to assign a possible origin of *anó* to the prefix *no-* and *asó* to the prefix *so-*.

Two other prefixes, *taga-* and *tagu-* were also examined and speculation was made as to their origins, respectively, from the word bases *tagá* and *tágo'*.

A number of infixes were also examined. Of these, *-al-* has obviously long been fossilised and it is not possible to assign a consistent meaning to it on the basis of evidence available. Discussion focussed on its relationship to the Bikol plural infix *-Vr-* and affixes of similar form in Tagalog and Malay.

The infix *-imin-*, obviously very much in evidence in the Bikol of the early seventeenth century, has completely disappeared from the modern language. Another affix of the form *-um-* functions differently in modern Bikol. What does exist in the modern form of the language is an infix of the form *-umin-* which seems to combine the somewhat stative function of both of these earlier affixes.

For practical as well as methodological reasons this paper has confined itself to the data available for one language, from one source, for one period of time. It is hoped, however, that the data presented will be of use to other linguists interested in the broader scope of Proto Austronesian reconstruction.

#### BIBLIOGRAPHY

- ASMAH HAJI OMAR, 1986, *Nahu Melayu Mutakhir*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- BLAKE, Frank R., 1925, *A grammar of the Tagalog language*. New Haven, Connecticut: American Oriental Society.
- BLUST, Robert A., 1980, Austronesian etymologies. *Oceanic Linguistics* 19:1-181.  
1984, Austronesian etymologies II. *Oceanic Linguistics* 22-23:29-149.  
1988, *Austronesian root theory: an essay on the limits of morphology*. Amsterdam/Philadelphia: John Benjamins.
- ENGLISH, Leo James, 1986, *Tagalog-English dictionary*. Manila: Capital Publishing House, for Congregation of the Most Holy Redeemer.
- GAN, Rosario Bella, 1964, *Balarilang pandalubhasaan*. Manila: Philippine Book Company.
- LISBOA, Marcos de, 1754, *Vocabulario de la lengua Bicol*. Sampaloc, Manila: Convento de Nuestra Señora de Loreto.
- MEES, C.A., 1969, *Tatabahasa dan tatakalamat*. (7th printing.) Kuala Lumpur: University of Malay Press.
- MINTZ, Malcolm Warren and José Del Rosario BRITANICO, 1985, *Bikol-English dictionary*. (With English-Bikol index.) *Dikسیونáring bikól-ínglés*. Manila: New Day Publishers.

The first study was a cross-sectional study of 1000 patients with rheumatoid arthritis who were treated with aspirin. The study found that patients who were treated with aspirin had a significantly higher rate of cardiovascular disease compared to those who were not treated with aspirin.

The second study was a longitudinal study of 1000 patients with rheumatoid arthritis who were treated with aspirin. The study found that patients who were treated with aspirin had a significantly higher rate of cardiovascular disease compared to those who were not treated with aspirin.

The third study was a cross-sectional study of 1000 patients with rheumatoid arthritis who were treated with aspirin. The study found that patients who were treated with aspirin had a significantly higher rate of cardiovascular disease compared to those who were not treated with aspirin.

The fourth study was a longitudinal study of 1000 patients with rheumatoid arthritis who were treated with aspirin. The study found that patients who were treated with aspirin had a significantly higher rate of cardiovascular disease compared to those who were not treated with aspirin.

### CONCLUSIONS

The results of these four studies suggest that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

These findings are consistent with the hypothesis that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

The results of these four studies suggest that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

These findings are consistent with the hypothesis that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

The results of these four studies suggest that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

These findings are consistent with the hypothesis that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

The results of these four studies suggest that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

These findings are consistent with the hypothesis that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

The results of these four studies suggest that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

These findings are consistent with the hypothesis that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

The results of these four studies suggest that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

These findings are consistent with the hypothesis that aspirin treatment may be associated with a higher rate of cardiovascular disease in patients with rheumatoid arthritis.

## THE SAMOAN CONSTRUCTION OF REALITY

ULRIKE MOSEL

### 1. INTRODUCTION

The following paper presents a short and very preliminary contrastive analysis of how simple statements about events or states of being are expressed in Samoan and English. The main difference between English and Samoan in this respect is that in English expressions referring to persons tend to occur in syntactically more prominent functions than expressions referring to objects or abstract phenomena, whereas in Samoan the expressions referring to persons are much more frequently found in syntactically subordinate functions. Therefore, English and Samoan can be characterised as person-oriented and anti-person-oriented languages respectively.

Our point of departure is the observation that somehow everything can be said in each of these two languages, but that it is almost impossible to translate from one language into the other without sacrificing some characteristics of the original or without using uncommon expressions in the translation. According to Grace (1981:177) a translation is an

expression in a second language of an idea which has been expressed in a first one. Translation...holds the idea constant without regard for different audience characteristics, and whether or not aesthetic form, sociolinguistic and expressive modulation, or content form of the original are represented accurately.

A literal translation preserves the content form of the original at the cost of the idiomaticity of the target language, whereas a free translation meets "the requirements of the idiomaticity of the target language", but "sacrifices the content form of the original" (Grace 1981:177).

These definitions suggest that the difficulties we have when translating from one language into the other result from one or more of the following factors:

- (a) different audience characteristics;
- (b) different aesthetic forms;
- (c) different sociolinguistic and expressive modulation;
- (d) different content forms.

The following analysis will concentrate on a very limited aspect of the content form, namely on certain syntactic structures and their semantic and pragmatic functions.

## 2. SYNTACTIC ASPECTS OF THE CONTENT FORM OF UTTERANCES IN ENGLISH

Before we turn to the syntactic, semantic and pragmatic features of Samoan which are difficult to translate, I would like to include here a few preliminary reflections on what it means to say that someone makes a statement about his or her experiences or observations. A central notion in Grace's definition of translation is that of content form. The content form of an utterance is "the way in which the idea which it expresses is analyzed (construed) for expression – the way it is put into words. This construction in fact creates a model of a bit of reality (or as- if reality)" (Grace 1981:177).

In order to understand how an idea is put into words, let us consider what happens when somebody observes a certain event – a transitive action, for instance – and then expresses what has been observed in a single sentence, such as:

- (1) 'The boy hit the dog'. or
- (2) 'The dog was hit by the boy'.

By definition a transitive action involves a person or animal which consciously initiates and performs the action in question, that is, the agent. Secondly, it involves the so-called patient – a person, animal, or object in the broadest sense – which is affected by the action, and thirdly, the place, duration and point of time of the action. What the reporter of such an event has observed is multidimensional. But when he puts his observation into words, he must analyse it into its components (more exactly, into what he considers its components), then select those details he wants to express, and finally, he must arrange what he selected for expression into a linear sequential order. Any stretch of speech or writing is linear.

Usually the speaker has more than one option for arranging the expression equivalents of what he conceptualises as components of the state of affairs in question. Thus, when reporting a transitive action like hitting, a speaker of English can choose between an active and a passive sentence construction, for example between (1) 'The boy hit the dog', and (2) 'The dog was hit by the boy'. Which of these constructions he actually chooses depends on whether he wants to make a predication about the agent or the patient. The linguistic utterance not only expresses a content but puts the content into a certain form. As far as simple clauses like (1) and (2) are concerned, the speaker of English has a choice between two perspectives of presentation. He can choose either the agent or the patient as the point of departure, as the subject or grammaticalised topic.

Whenever a speaker of English makes a statement about a state of affairs and uses an independent sentence as a means of expression, he has to structure the information into a subject and a predicate. And if there is actually no entity about which a predication can be made, he will employ the pronoun 'it' as a dummy subject, for example, 'It is raining'. Since the subject is such a central category of English syntax, it seems worthwhile with regard to our contrastive analysis of English and Samoan to have a look at some of its syntactic properties (Keenan 1976):

### (a) Grammatical indispensability

All syntactically independent clauses must have a subject in English; without a subject they would be ungrammatical. One cannot say \*'Hit the dog', but rather 'The dog was hit', if one wishes to omit the agent.

### (b) Word order

The unmarked word order in declarative sentences is subject + verb (+ object).

## (c) Agreement between subject and predicate

If there is agreement in number between an argument and the predicate, it is the subject which agrees in number with the predicate and controls the agreement: 'The boy is chasing the dogs. The dogs are being chased by the boy'.

## (d) Coreferential deletion across conjunctions

If one coordinates two predications about the same subject, it is unnecessary to express the subject twice: 'The boy hit the dog and (the boy/he) ran away'. Or to put it differently, the sentence 'The boy hit the dog and ran away' can only mean that the boy ran away.

## (e) Subject = agent in active transitive clauses

The structurally simplest clause to express a transitive action is the active clause, in which the subject signifies the agent.

To conclude, in English, statements are always structured as statements about a person, an animal, an object or a phenomenon. If a person is involved in some state of affairs, it is more likely that the speaker will form the statement as predication about that person than that he will choose any other construction in which the person plays a less salient role. This is particularly true in statements about actions and mental or bodily states of being or activities. Instead of the unmarked active clause 'The boy hit the dog' the more marked passive clause 'The dog was hit by the boy' is only used if the dog is explicitly the topic of discourse. If, however, in a statement about some affectedness, the affected participant is a person and what affects him or her is something inanimate, the passive construction seems to be more likely:

(3) 'The boy was hit by a stone'.

Another way of expressing such a state of affairs is to employ a patient-oriented active verb whose subject denotes a patient:

(4) 'The boy suffered from a headache'.

'The boy has got the measles'.

In other words, the English ways of expression focus on the person, if any is involved, as they assign a subject-predicate structure to the statement and select the noun phrase referring to the person for the grammatical subject. There is only one way to avoid subject-predicate constructions when making statements, and that is nominalisation. However, nominalisations are not used as syntactically independent clauses, but are embedded in other clauses where they function as complements or adverbials. Furthermore, their use is idiomatically restricted to certain, mostly formal, contexts.

### 3. BASIC CHARACTERISTICS OF SAMOAN CLAUSE STRUCTURE

In contrast to English, Samoan statements about events or states of being do not show a subject-predicate structure when they are expressed by basic verbal clauses. Samoan lacks the grammatical categories of subject and object, and does not distinguish between active and passive voice or ergative and antipassive constructions. Consider the following clauses which show how states of being and



intransitive and transitive actions are expressed in Samoan (there are lots of other types of states of affairs expressed by the same or different constructions):<sup>1</sup>

- (5) 'Ua ma'i le tama.<sup>2</sup>  
 PERF sick ART child  
 verb phrase absolutive noun phrase  
 The child is sick.
- (6) 'Ua alu le tama.  
 PERF go ART child  
 verb phrase absolutive noun phrase  
 The child has gone.
- (7) Na fasi e le tama le maile.  
 PAST hit ERG ART child ART dog  
 verb phrase ergative noun phrase absolutive noun phrase  
 The child hit the dog.

In these clauses the arguments of the verb phrase are expressed by two kinds of noun phrases. Firstly, the argument of the intransitive clauses (5) and (6) and the argument denoting the patient in (7) are expressed by unmarked noun phrases, the so-called absolutive noun phrases. Secondly, the argument denoting the agent of a transitive action as in (7) is expressed by an ergative noun phrase, that is, a noun phrase marked by the ergative preposition *e*.

Some linguists who think that every language in the world must have a subject like English describe this Samoan case-marking pattern approximately as follows: 'The subject of intransitive clauses and the object of transitive clauses are expressed by noun phrases in the absolutive case, whereas the subject of transitive clauses is expressed by a noun phrase in the ergative case' (compare Chung 1978:54f.; Cook 1988:29f.; Ochs 1988:58f.). A closer look at Samoan grammar, however, reveals that these 'subjects' lack some important subject properties apart from the fact that they – in contrast to English subjects – show distinctive codings.

If there is agreement in number between the verb and an argument, it is the absolutive noun phrase in intransitive and transitive clauses which controls agreement in number:

- (8) 'Ua ma-ma'i tamaiti.  
 PERF PL-sick child(SP.PL)  
 verb phrase absolutive noun phrase  
 The children are sick.

<sup>1</sup>Abbreviations used are as follows: 1,2,3 – first, second, third person, ANAPH – anaphoric pronoun, ART – article (when not specified = specific singular *le*), CAUS – causative, CONJ – conjunction, DIR – directive particle, EMPH – emphatic particle, ERG – ergative (*e*), GENR – general tense-aspect-mood particle, LD – locative-directional (*i*, '*i*), NR – nominaliser, NSP – non-specific, OPT – optative (*se i*), PERF – perfect, PL – plural, POSS – possessive (*a* and *o*), PRES – presentative ('*o*), Q – question particle, SG – singular, SP – specific, TR – transitive.

<sup>2</sup>Samoan sentences which were recorded by myself are rendered in the orthography used by Milner (1966), whereas all examples taken from published texts are quoted in their original orthography, which often neglects the indication of glottal stops and vowel length.

- (9) 'Ua  $\bar{o}$ <sup>3</sup> tamaiti.  
 PERF go(PL) child(SP.PL)  
 verb phrase absolutive noun phrase  
 The children have gone.
- (10) Na fasi e tamaiti le maile.  
 PERF hit ERG child(SP.PL) ART dog  
 verb phrase ergative noun phrase absolutive noun phrase  
 The children hit the dog.
- (11) Na fa-fasi e le tama maile.  
 PAST PL-hit ERG ART child dog(SP.PL)  
 verb phrase ergative noun phrase absolutive noun phrase  
 The child hit the dogs.

These coding properties of the predicate and its arguments – that is, case marking and number agreement – argue against the claim that the ergative noun phrase functions as the subject. Furthermore, two other properties of absolutive and ergative arguments do not support this claim. As far as the so-called rule of coreferential deletion is concerned, Samoan has a very simple rule: whenever it is clear from the context which person, animal or object is involved in an action as the agent or patient, that participant does not need to be expressed:

- (12) Tū atu loa lea 'o Sina tago 'i le lupe,  
 stand DIR immediately then PRES Sina take.hold.of LD ART pigeon  
 titina, togi 'i fafo.  
 strangle throw LD outside  
 Sina stood up, took hold of the pigeon, strangled it, then threw it outside.  
 (Moyle 1981:102)

The most literal translations of 'The child hit the dog and ran away' and 'The dog was hit by the child and ran away' are, respectively:

- (13) Sā sasa e le tama le maile, ona sola ai lea 'o ia.  
 PAST hit ERG ART child ART dog CONJ run.away ANAPH that PRES 3SG  
 The child hit the dog, then he ran away.

and

- (14) Sā sasa e le tama le maile, ona sola ai lea  
 PAST hit ERG ART child ART dog CONJ run.away ANAPH that  
 le maile.  
 ART dog  
 The child hit the dog, then the dog ran away.

In (13) *sola* is combined with 'o ia which unambiguously refers to *le tama*, because 'o ia is only used to refer to human beings. Therefore, *le maile* has to be repeated in (14). In both cases *ona sola ai lea* '(and) then ran away' would not make sense, because the context provides no clue as to who ran away. Note that many informants who were asked in Samoan to combine the two clauses, *Sā sasa e le tama le maile* 'The child hit the dog' and *Sā sola le tama* 'The boy ran away' said:

<sup>3</sup>  $\bar{o}$  is the suppletive plural form of *alu* (SG) 'go'.

- (15) *Sā sola le tama ina 'ua sasa le maile.*  
 PAST run.away ART child CONJ PERF hit ART dog  
 The child ran away when he had hit the dog.

Correspondingly, *Sā sasa e le tama le maile* 'The child hit the dog' and *Sā sola le maile* 'The dog ran away' were combined as:

- (16) *Sā sola le maile ina 'ua sasa e le tama.*  
 PAST run.away ART dog CONJ PERF hit ERG ART child  
 The dog ran away when it had been hit by the child.

To conclude, the rule of coreferential deletion across conjunctions is a semantically determined rule, which argues for the absence of a grammaticalised topic and correlates with the absence of an active/passive or ergative/antipassive distinction. Moreover, in Samoan it is possible to make statements about successive intransitive and transitive actions without mentioning the agent.

- (17) [On Friday evening work was finished, the usual time for knocking off was long overdue. Overtime had to be done the whole week, and so the feeling finally arose that it was absolutely necessary to have a big party in the evening.]

*Se'i hepi tasi fo'i. 'Ua fia tau fo'i i le ua,*  
 OPT happy once again PERF want reach again LD ART neck  
*e fa'a-galo ai le tigaina.*  
 GENR CAUS-forget ANAPH ART pain

To be happy once again. To get one's fill in order to make the pain be forgotten. [This longing for freedom coincided perfectly with an invitation that had been received for just this night. The most important thing was drinking for nothing.]

*Na ona<sup>4</sup> taunu'u lava i le fale, fa'amalie tamaiti*  
 only arrive EMPH LD ART house put.at.ease child(SP.PL)

*ma le to'alu i upu, ma le polomisi e vave mai*  
 and ART wife LD word(SP.PL) with ART promise GENR soon DIR

*i le fale, nimo loa. Na fo'i i le fale, 'ua te'a le*  
 LD ART house disappear then PAST return LD ART house PERF past ART  
*tasi i le vaveao...*  
 one LD ART dawn

Came home, put the kids and wife at ease with the promise to return home soon, then disappearing. Returned home, it was already past one in the morning... (Sootaga 1981:1)

- (18) *'Ua ola le taavale ma alu ese. 'Ua tatala le faitotoa tele*  
 PERF life ART car and go away PERF open ART door big  
*ona savavali mai lea i totonu.*  
 CONJ walk(PL) DIR that LD inside  
 [The narrator who is in prison reports:]  
 The car started and went away. The big door opened. Then some people<sup>5</sup> went inside.  
 (Tuitolovaa 1983:1)

<sup>4</sup> *Na ona* has to be analysed as *na* 'only' + *ona* (CONJ) 'that'.

<sup>5</sup> Neither in the preceding nor in the following context are the people who went inside explicitly mentioned.

In Samoan the content form of statements about actions does not require the expression of the agent, whether the actions are transitive or not. In respect to the indispensability of the so-called 'subject' this would mean that subjects are dispensable if they denote the agent of an intransitive or transitive action. This would hardly make sense, as it excludes another subject property, so that in the end nothing is left but the English translations of clauses like (5)–(7), which would argue for the existence of the grammatical category of subject. However, Samoans do have the possibility of presenting a statement about a state of affairs as a predication about one of the participants. But then they cannot use the simple verbal clause, but have to use a more complex type of construction:

(19) 'O le tama na fasi-a le maile.  
 PRES ART boy PAST hit-TR ART dog  
 The boy hit the dog. / It was the boy who hit the dog.

(20) 'O le maile na fasi e le tama.  
 PRES ART dog PAST hit ERG ART boy  
 The dog was hit by the boy. / It was the dog which was hit by the boy.

The noun phrase denoting the participant about whom the predication is made is marked by 'o (the preposition signifying the presentative case) and placed at the beginning of the sentence. If this fronted noun phrase refers to the agent of a transitive action (19), the verb denoting the action is usually marked by the so-called 'mysterious transitive suffix' (Cook 1978; Mosel 1985). Compared with the basic verbal clauses, the clauses with fronted noun phrases are a marked construction (Mosel 1987).

#### 4. ANTI-PERSON-ORIENTED FORMS OF SAMOAN UTTERANCES

Examples (5)–(7) suggest that, in spite of the absence of the syntactic categories subject and object, the Samoan absolutive and ergative noun phrases correspond semantically to English subject and object noun phrases. However, they do so only to a certain degree. It is not the case that every English transitive clause would be translated by a Samoan verbal clause with an absolutive and an ergative noun phrase.

As mentioned above, English ways of expression focus on the person spoken about and tend to refer to him/her by the subject noun phrase. But in Samoan there is a clear tendency to avoid expressions which foreground the person. Consider examples (21) and (22), which illustrate the following rule: if the agent of a transitive action is the possessor of the patient, it is not expressed by an ergative noun phrase or a preverbal pronoun, but by a possessor noun phrase or possessive pronoun which syntactically functions as an attribute of the patient noun phrase (Duranti 1981:173; Duranti and Ochs 1990; Mosel 1985:102f., 1989).

(21) 'O le ā l-a-'u mea na 'ai?  
 PRES ART what ART-POSS-2SG thing PAST eat  
 Na 'ai l-a-'u fasi talo.  
 PAST eat ART-POSS-1SG piece taro  
 What did you eat? I ate a piece of taro.  
 (lit. What is your thing you ate? My piece of taro was eaten.)

- (22)
- Na luelue le ulu o Iopu.*

PAST shake ART head POSS Iopu.

**Iopu** shook his head. (lit. Iopu's head was shaken.) (Taulogo 1986:2)

When Samoans speak of perceptions, emotions or feelings, they prefer expressions in which the pronoun or noun phrase denoting the experiencer functions as a possessive attribute. It is either the attribute of a nominalised verbal construction denoting the experience (23), (24) or the attribute of a noun phrase referring to a body part which is somehow involved in the experience (25), (26); compare Shore 1982:172f.

- (23)
- Na lagona le faanoanoa o le taule'ale'a o Lama ina ua*
- 
- PAST feel ART sad POSS ART untitled.man PRES Lama CONJ PERF

*tau atu l-a-na vaa i l-a-na teine.*

reach DIR ART-POSS-3SG see LD ART-POSS-3SG girl

**The untitled man** Lama felt sad when **he** caught sight of his girl.

(lit. The being sad of the untitled man Lama was felt when his seeing reached his girl.)

(Samoa Weekly, Sept. 1989:2)

- (24)
- 'Ua faanoanoa o-u lagona.*

PERF sad POSS-1SG feel

**I** am sad. (lit. My feelings are sad.) (Leiataua 1986:10)

- (25)
- 'Ua tete l-o-ta moa.*

PERF tremble ART-POSS-1SG solar plexus

**I** am scared. (lit. My solar plexus trembles.) (Milner 1966:146)

- (26)
- 'Ua tutū moge o l-o-na tino.*

PERF stand(PL) hair POSS ART-POSS-3SG body

**He** had goose flesh. (lit. The hairs of his body stood erect.) (Milner 1966:147)

- (27)
- Oo mai loa Tulua i le fale, faataga tepa ese mata*
- 
- come DIR then Tulua LD ART house pretend look away eye(SP.PL)

*o le tina a o taliga la e*  
POSS ART moth but PRES ear(SP.PL) EMPH GENRer*faalogologo ma maitau soo se pao.*

hear and notice every noise

When Tulua came home, the eyes of her mother pretended to look away, but her ears listened and noticed every noise. (Semau 1979:4)

Similarly, the appropriate expression to ask somebody for a favour is *E malie lou loto...* (or more politely *E malie lou finagalo...*) 'Does your heart/will agree...?', which corresponds to English 'please', but in contrast to 'please', it does not directly address the hearer:

- (28)
- E malie l-o-u loto pē mafai ona ou fa'aaogā-ina*
- 
- GENR agree ART-POSS-2SG heart Q possible CONJ 1SG use-TR

*l-a-u penitala.*

ART-POSS-2SG pencil

Could I, please, use your pencil? (lit. Does your heart/will agree, if it is possible that I use your pencil ?)

Statements about a person's age follow the same pattern. While a numeral forms the predicate, the pronoun or noun phrase referring to the person is encoded as a possessive attribute of *tausaga* 'years' which functions as the primary argument of the predicate:

- (29) *E ono tausaga o le teine.*  
 GNR six year(SP.PL) POSS ART girl  
 The girl is six years old. (lit. The years of the girl are six.)

A further feature which distinguishes Samoan from English is its extensive use of nominalisations in which a verb or a verbal noun derived by =*ga* forms the nucleus of a noun phrase. The participants in the state of affairs denoted by the verb or verbal noun take the form of attributes; in other words, they are given the form of a subordinated constituent. The following examples illustrate three typical usages of nominalisations which do not correspond to English expressions containing nominalisations. Apart from these three types of nominalisations which cannot be translated literally into English, Samoan has all usages found in English:

(a) In narrative texts, nominalisations often occur in the presentative case at the beginning of complex sentences, where they describe a certain state of affairs as the background to what the following sentence is about:

- (30) *O le tago ifo o Iopu i l-a-na afitusi*  
 PRES ART reach down POSS Iopu LD ART-POSS-3SG match  
*e tutu ai l-a-na muli-tapaa*  
 GNR light ANAPH ART-POSS-3SG end-tobacco  
*sa sei i l-o-na taliga*  
 PAST put.behind.the.ear LD ART-POSS-3SG ear  
*ae tau-valaau mai loa Maatusi...*  
 but repeatedly-call DIR then Maatusi  
 Iopu was just reaching down for his match to light his cigarette end which he had put behind his ear when Maatusi repeatedly called him... (lit. The reaching down of Iopu for his match...) (Taulogo 1986:1)

(b) Nominalised verbal clauses form exclamative nominal clauses which are a very common means of enthusiastically commenting on a certain situation (cf. Duranti and Ochs 1990):

- (31) *O le manaia ia o l-a-u moe i lena po.*  
 PRES ART nice EMPH POSS ART-POSS-1SG sleep LD that night  
 I had a wonderful sleep that night! (lit. The being nice of my sleeping in that night!)  
 (Larkin 1967:7)

(c) Verbs of motion like *alu* 'to go', *fo'i* 'to return', *sau* 'to come', and *taunu'u* 'to arrive' can be combined with nominalised verbal clauses to express the fact that somebody is going to do something or is coming from having done something:

- (32) *Sa alu loa le utu-ga-sami a le tamaitai.*  
 PAST go then ART scoop-NR-salt-water POSS ART lady  
 The girl went to get salt-water. (lit. The salt-water scooping of the lady went.)  
 (Sio 1984:13)



- (33) *'Ua sau le tā'ele= ga a le teine.*  
 PERF come ART bathing=NR POSS ART girl  
 The girl has come from having a bath.

The last kind of anti-person-oriented expressions to be mentioned here are possessive constructions. While in English possession is expressed by clauses of the type 'X has (owns, possesses) Y', in which the possessor X functions as the subject, Samoan forms possessive clauses with the existential verbs *iai* 'exist' and *leai* 'not exist, be absent', in which the possessee Y forms the primary argument of the existential predicate and the possessor X an attribute of this argument. The literal translations of the Samoan construction would be 'The Y of X exists/does not exist'. Samoan lacks verbs meaning 'to have, own, possess'.

- (34) *E iai le ilāmutu o Saētānē 'o le ve'a...*  
 GNR exist ART aunt POSS Saētānē PRES ART rail  
 Saētānē had an aunt who was a rail... (Moyle 1981:50)
- (35) *E tatau ona iai s-a-u laisene.*  
 GNR necessary CONJ exist ART(NSP.SG)-POSS-2SG licence  
 You need a licence. (lit. It is necessary that your licence exists.) (Milner 1966:95)
- (36) *E iai l-a-'u telefoni.*  
 GNR exist ART-POSS-1SG telephone  
 I have a telephone.
- (37) *E leai s-a-'u tupe.*  
 GNR not.exist ART(NSP.SG)-POSS-1SG money  
 I don't have money.

## 5. CONCLUSION

As the preceding examples illustrate, Samoan differs considerably from English in how real-world situations are put into words.

(a) In their basic forms, expressions of events or states of being in Samoan are not designed as statements of a subject-predicate structure. Instead, the event or state as such is denoted by a clause-initial verb phrase which is followed by noun phrases giving additional information about who or what was involved. This information may be missing in many contexts where, for grammatical reasons, the English translation equivalent must explicitly refer to the participants.

(b) Samoan makes extensive use of nominalisations in which, to an even greater degree than in basic verbal clauses, the expression of the situation as such forms the core of the linguistic construction, whereas all participants are denoted by optional attributes and given a subordinate status.

(c) In statements about a person's experiences, age or belongings, the person is signified by a subordinate possessor phrase and is thus, linguistically, put in the background.

The linguistic differences between Samoan and English doubtlessly correspond to cultural differences. But for the time being it is very difficult, if not impossible, to say to what degree linguistic phenomena directly reflect the sociocultural background of the speakers of either language, or whether or not a certain way of expression says anything about a person's way of thinking. It is

"impossible to draw a clear line between thinking, i.e. bringing a thought into being, and encoding the thought, i.e. putting it into words" (Grace 1987:10). Therefore, the present article does not make any statements about the interrelationship between language and culture, or language and thinking, but only seeks to show some essential differences between English and Samoan ways of expression.

## BIBLIOGRAPHY

- AIOVAO, Tunumafono Apelu, 1987, *Maunu mai loimata o Apa'ula*. Apia: Institute of Pacific Studies and the Western Samoan Extension Centre of the University of the South Pacific and the Inivesite Aoa o Samoa (National University of Samoa).
- CHUNG, Sandra, 1978, *Case marking and grammatical relations in Polynesian*. Austin: University of Texas Press.
- COOK, Kenneth, 1978, The mysterious Samoan transitive suffix. In Jeri Jaeger, Anthony C. Woodbury et al., eds *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistic Society*, 53-66. Berkeley, Cal.: Berkeley Linguistic Society.
- COOK, Kenneth, 1988, A cognitive analysis of grammatical relations, case, and transitivity in Samoan. PhD dissertation, University of California, San Diego.
- DURANTI, Alessandro, 1981, *The Samoan fono: a sociolinguistic study*. PL, B-80.
- DURANTI, Alessandro and Elinor OCHS, 1990, Genitive construction and agency in Samoan discourse. *Studies in Language* 14/1:1-23.
- GRACE, George W., 1981, *An essay on language*. Columbia, South Carolina: Hornbeam Press.
- 1987, *The linguistic construction of reality*. London, New York, Sydney: Croom Helm.
- KEENAN, Edward, 1976, Towards a universal definition of "subject". In Charles Li, ed. *Subject and topic*, 305-333, New York: Academic Press.
- LARKIN, Fanaafi, 1967, *O le tala a Mandy Jane*. Apia: Ofisa e saunia tusi aoga, matagaluega o aoaoga.
- LEIATAUA, Iosefa Iosefa, 1986, Pale le maumea, ula foi le faasususu. *Moana* 1986:10-14.
- MILNER, George B., 1966, *Samoan dictionary*. London: Oxford University Press.
- Moana: O se tusi i le gagana Samoa*. (Vol.1 1979 28 pp., Vol.2 1980 28pp., Vol.3 1981 32pp., Vol.4 1982 28pp., October 1983 28pp., Oketopa 1984 44pp., Oketopa 1985 45pp., Oketopa 1986 44pp., Oketopa 1987 44pp., Novema 1987 44pp.) Apia. U.S.P. Centre.
- MOSEL, Ulrike, 1985, *Ergativity in Samoan*. Arbeiten des Kölner Universalien Projekts 61. Köln: Institut für Sprachwissenschaft.
- 1987, Subject in Samoan. In Donald C. Laycock and Werner Winter, eds *A world of language: papers presented to Professor S.A. Wurm on his 65th birthday*, 455-479. PL C-100.
- 1989, On the classification of verbs and verbal clauses in Samoan. In Ray Harlow and Robin Hooper, eds *VICAL 1: Papers from the Fifth International Conference on Austronesian Linguistics*, 377-398. Auckland: Linguistic Society of New Zealand, 377-398
- MOYLE, Richard, 1981, *Fāgogo: fables from Samoa in Samoan and English*. Auckland: Auckland University Press.
- OCHS, Elinor, 1988, *Culture and language development: language acquisition and language socialization in a Samoan village*. Cambridge: Cambridge University Press.
- SEMAU, Faamau, 1979, Aoga Afiafi. *Moana* 1979:4-5.
- SHORE, Bradd, 1982, *Sala'ilua: a Samoan mystery*. New York: Columbia University Press.
- SIO, Gaoloaifaana Peseta, 1984, *Tapasā o folauga i aso afā. Compass of sailing in the storm*. Apia: U.S.P. Centre.
- SOOTAGA, Vaea, 1981, Si alii ma'i. *Moana* 1981:1-4.
- TAULOGO, Fiapopo, 1986, Pomu Niukilia. *Moana* 1986:1-4.
- TUITOLOVAA, Agafili Laau, 1983, E fai taliga o aupa. *Moana* 1983: 1-4.

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: (773) 936-3000

THE UNIVERSITY OF CHICAGO LIBRARY

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: (773) 936-3000

THE UNIVERSITY OF CHICAGO LIBRARY

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: (773) 936-3000

THE UNIVERSITY OF CHICAGO LIBRARY

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: (773) 936-3000

THE UNIVERSITY OF CHICAGO LIBRARY

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: (773) 936-3000

THE UNIVERSITY OF CHICAGO LIBRARY

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: (773) 936-3000

THE UNIVERSITY OF CHICAGO LIBRARY

THE UNIVERSITY OF CHICAGO LIBRARY  
1215 EAST 58TH STREET  
CHICAGO, ILLINOIS 60637  
TEL: (773) 936-3000

RELATIONAL MORPHEMES AND A TRANSITIVISING SUFFIX IN XÂRÂCÙÙ  
(NEW CALEDONIA)

CLAIRE MOYSE-FAURIE

1. INTRODUCTION

Present day Xârâcùù<sup>1</sup> has only one transitive verb suffix (-*ri*), which can be applied to only a few verbs. The loss of Proto Oceanic transitive suffixes has been compensated for by the development of several relational morphemes or case markers. This paper discusses the origin and use of the latter in relation to verb serialisation.

Xârâcùù is a serialising language in which verb series tend to give rise:

- either to grammatical morphemes (relational morphemes or modal auxiliaries) or postverbs resulting from specialisation of one of the two verbs in the series ;
- or to transitive compound verbs, generally obtained by shortening the syllable structure of the first verb in the series.

The argument markers which introduce the objects of otherwise intransitive verbs can (with the exception of *wã* 'for, about', cf. 6.2) be placed fairly at will within the verb phrase or before the object they introduce.

Diachronically speaking, some markers of participant relationships derive:

- from former postverbs which become object markers (*ce* 'towards, for', cf. 3.)
- from verbs which become grammaticalised in series (*xi* 'give', *taa* < *witaa* 'take away', *tara* 'see', cf. 2.4).
- from prepositions which tend to become verb-enclitic when they introduce an object (*ti* 'at, on', *ŋɛ* 'with', cf. 5. and 6.1).

This paper will deal with three main points. The first will be the facts of verb serialisation and how it can develop into verb + adverb, auxiliary + verb, or verb + marker sequences. The behaviour of the relational markers introducing the objects of verbs will be described in terms of their choice of positions, their origin, and the classes of verbs they associate with.

---

<sup>1</sup>Xârâcùù has about 4000 speakers in the Thio-Canala region of New Caledonia.

Robert Blust, ed. *Currents in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, 305-320. *Pacific Linguistics*, C-117, 1991.

© Claire Moyse-Faurie

This will be followed by a brief discussion of verb compounding. This process, which is richly productive in the southern Caledonian languages, is likely to result in part from serialisation, and affects verb valency.

Finally, the range of diversity in these markers will be brought out by way of short comparisons with neighbouring languages. This diversity is a severe hindrance to syntactic comparison and reconstruction in the southern Caledonian languages. The difficulties are comparable to those encountered by G.W. Grace in his attempted phonological reconstructions.

## 2. VERB SERIALISATION

The verb phrase in Xârâcùù is generally composed of a verb or verb series with following or preceding adverbs, bounded by various tense-aspect markers. The lexical or pronominal subject is, like the object, outside the limits of the verb phrase, which need contain no mark of grammatical person.

Verb serialisation consists of a sequence of two verbs, each of which could appear independently. Serialisation, unlike compounding, involves no formal changes. Verb series make up an indivisible whole and contain no relational morphemes. They share a common case frame, common modifiers, and common tense-aspect markers. As a whole, they have a special falling accent pattern: the main accent falls on the first syllable of the first verb, with an accent on a lower register on the first syllable of the second verb.<sup>2</sup> One of the two verbs in the series often tends to specialise in a modal, adverbial or relational role.

### 2.1 SEQUENCE OF TWO VERBS

The semantics of a verb series will generally combine a position and a movement, two positions, or an action and the way of carrying it out or its intended purpose. Here are a few examples:

<i>kwe pwāāī</i>	<i>ʃʌga</i>	<i>cue</i>
dance turn	warm in the sun	sit
dance spinning about	get warm sitting in the sun	
<i>kē façaa</i>	<i>pūxūrū tēī</i>	
eat spoil	run cry	
waste food in the process of eating	run crying	
<i>ʃe çāā</i>	<i>nɔu</i>	<i>xwīī</i>
pull stretch	cut (sugar cane)	chew
pull for the purpose of stretching	cut sugar cane to chew on it	

### 2.2 SEQUENCE OF TWO VERBS (THE FIRST SPECIALISES)

One of the two verbs tends to appear more frequently in serialisation than independently, and becomes specialised in a modal or adverbial role.<sup>3</sup>

<sup>2</sup>For a description of accent in Xârâcùù, see Rivierre (1978).

<sup>3</sup>Abbreviations used are as follows: ANA – anaphora, COLL – collective, DU – dual, DUR – durative, E – exclusive, I – inclusive, ING – ingressive, NEG – negative, NOM – nominalising prefix, PL – plural, POSS – possessive, PROHIB –

**modal specialisation:**

*nããbu* 'begin', as in:

*xiti wã nããbu*  
celebration ING begin  
The celebration begins.

becomes a modal auxiliary when followed by another verb:

*ε nããbu ïnũ kεε-fãrã*  
3SG begin learn NOM-read  
He begins to learn to read.

Likewise, *nãã* 'ask for':

*nã nãã taa ro çãã-pwĩ*  
1SG ask for REL 2SG sucker-banana tree  
I ask you for banana-tree suckers.

*nãã* acts as a modal auxiliary meaning 'want' when it precedes another verb:

*a xũũ/ĩ a nãã metiçe*  
this child this ask for sleep  
This child wants to sleep.

**adverbial specialisation:**

*tete* 'wander about' rarely appears separately:

*dεeri nã tete*  
people DUR wander  
The people walk about.

It is usually proposed to verbs of movement with the meaning 'to no purpose, for no reason':

*ĩrĩ nã tete ïεe*  
1PL.I DUR wander go down  
We go down for no particular reason.

**2.3 SEQUENCE OF TWO VERBS (THE SECOND SPECIALISES)**

Some second verbs in a series tend to specialise not only in modal or adverbial senses but also as relational morphemes, and then break off from the verb phrase; *verb + verb + object sequence* then becomes *verb + relational morpheme + object*.

**modal specialisation**

*wĩnã* 'pursue (something)':

*ε wĩnã wake rεε*  
3SG pursue work POSS.REL.3SG  
He continues his work.



The verb *wĩnĩ* is more often used modally with the meaning 'keep on':

*ε da wĩnĩ*  
3SG eat pursue  
He keeps on eating.

### adverbial specialisation

*xoru* 'good, beautiful', a stative verb as in:

*a kãmũrũ a xoru*  
this man this handsome  
This man is handsome.

means 'well, nicely' after another verb:

*ε kwe xoru*  
3SG dance well  
He dances well, is a good dancer.

### adverbial or relational specialisation:

*jati* 'work especially for someone':

*ŋěě jati laake xĩja ẽẽ xi-tia*  
1PL.E work for many bracelets RES scratch-split  
We work intentionally to make precut bracelets.

After another verb, *jati* marks the recipient of the action:

*xũ jati* 'plough for someone'

*pa kwara xũ jati aaxa re ri*  
COLL subject plough for chief POSS.REL 3PL  
The subjects plough for their chief.

*jai* 'pass beyond, transgress'

*ε jai dõu mwiri wei mwiri aaxa le na*  
3SG pass thing there this one there chief say PAST  
He does not respect what the chief said.

As the second verb in a series, *jai* will express the comparative:

*ε kaxẽ jai nã*  
3SG big pass 1SG  
He is taller than I.

*jati* and *jai*, are transitive verbs which seem about to become relational morphemes, but can never be separated from the preceding verb. On the other hand, *fadi* 'measure, reckon, make equitable distribution' can be removed from the verb phrase and used to introduce circumstantial modifiers.

As a separate verb:

*ri nĩ fadi re kɛɛ-mwaa re dõu mwiri ti xwãnə-ɾe /ẽẽde*  
3PL DUR reckon DUR NOM-long POSS.REL thing there to end-3SG evening  
They reckon that it will last until evening.

In a verb series:

*pa xũũf̃ fɫdɔ fadi çere-ri*  
 COLL child walk share way-3PL  
 The children have the same way of walking.

As a relational morpheme meaning 'according to', introducing a circumstantial modifier:

*ĩĩ xɫɫi pa xũũf̃ fadi dɔ tɛpə rɛ xũã*  
 1PL.I raise COLL child according to true custom POSS.REL tribe  
 We raise our children according to the ancient custom of the tribe.

## 2.4 RELATIONAL MORPHEMES DERIVED FROM SERIALISATION AND USED TO INTRODUCE OBJECTS OF VERBS

Some of the relational morphemes introducing indirect objects have no other function. Others are also postverbs or are still verbs in their own right, and can form compounds. They have varying degrees of autonomy with respect to the verb they transitive and the object they introduce.

Each of these relational morphemes is compatible with a small set of verbs, usually taken from a common semantic field (verbs of communication, verbs of feeling, pejorative verbs, verbs describing exchanges etc.)

### 2.4.1 RELATIONAL MORPHEMES DERIVED FROM STILL AUTONOMOUS VERBS

#### 2.4.1.1 *tara* 'know, see'

*tara* is now rarely used independently:

*kə tara, kwie çə-xuru rũ kɛ xũ kwãã a*  
 2SG see rain hit-flee 1DU.I from on trœ this  
 You see, the rain drives us out of the tree.

After another verb, *tara* may have an adverbial function and mean 'to see, for curiosity's sake'. In this role, *tara* immediately follows the verb it modifies:

*kə fã tara nenuu nœ*  
 2SG undo see package this  
 You undo this package for the sake of curiosity.

*ɛ kɛ tara m̃r̃õ na ku*  
 3SG eat see already PAST yam  
 He has already tasted the dish of yams.

Finally, *tara* seems to have acquired the role of a relational morpheme introducing the objects of certain verbs which describe attempts at communicating with another being: *jumĩ* 'make a sound of kissing'; *piaxõ* 'whistle'; *piçinõ* 'glance at'; *çĩĩ* 'make a noise to attract attention'. In this use, it can be separated from the first verb by aspect markers or postverbs and preposed to the object. Its meaning is then 'towards, at, for':

*nā piaxō mwāmwaa na tara deeri*  
 1SG whistle long PAST REL people  
 I whistled at the people (to attract their attention) for a long time.

2.4.1.2 *xi* 'give'

*xi* 'give' is a trivalent verb and is compatible with the relational morpheme meaning 'to (attributive)', to which it has itself given rise. This marker can be placed either within the verb phrase immediately after the verb, or immediately before the attributive object it introduces:

*kə xi xi na nū laa mwanɔ (=kə xi na laa mwanɔ xi nā)*  
 2SG give REL PAST 1SG a cloth  
 You gave me a piece of cloth.

The verbs which take the relational morpheme *xi* are verbs of communication or exchange. They invariably assume something being given away, an outward movement. The recipient is always an animate being. *xi* may introduce objects of otherwise intransitive verbs: *baa* 'show oneself, appear'; *nīmō* 'tell a story'; *ḡāā* 'cry out'; *xoru* 'please'; *yaaru* 'set riddles'; *xa* 'speak'; *le* 'say'; or attributive objects of transitive verbs: *nūmārā* 'give away'; *çu* 'write'; *çue* 'entrust'; *xipwei* 'announce'; *nēge* 'request respectfully'; *xacie* 'show'; *nū* 'send'; *xādii* 'pay'.

*aḡāā xa xi xūūfī a*  
 mummy speak REL child this  
 Mummy speaks to this child.

*nā xādii pwī xi aḡāā*  
 1SG pay bananas REL mummy  
 I pay Mummy for the bananas.

2.4.2 A RELATIONAL MORPHEME DERIVED FROM A BOUND VERB: *taa*

*taa* (< *witaa* 'throw away, take off') invariably retains a meaning of withdrawing or taking away. As a postverb, it means 'once and for all', 'to get rid of (something)'; in composition, it means 'leave', and as a relational morpheme, it means 'away from (DIS(associative))'.

***taa* as postverb with a modal meaning**

As a postverb, *taa* immediately follows the first (generally transitive) verb and precedes the aspect markers, thus behaving like the second verb in a verb series, although it never appears as a verb in its own right.

Compare utterances (a) and (b) below. The presence of *taa* gives the (b) utterances a terminative sense:

- (a) *wījō laa kwe mīā*  
 drink a water red  
 Drink the wine.
- (b) *wījō taa laa kwe mīā*  
 drink a water red  
 Drink all the wine, finish the bottle!

- (a) *ε fe tē re do-çaa*  
3SG go throw POSS.REL thing-bad  
He is going to throw the trash away.
- (b) *ε fe tē taa re do-çaa*  
3SG go throw POSS.REL thing-bad  
He is going to throw the trash far away (to be rid of it once and for all).
- (a) *pāādo ba na çñ re ri*  
men plough PAST hole POSS.REL 3PL  
The men made holes for their yams.
- (b) *pāādo ba taa na çñ re ri*  
men plough PAST hole POSS.REL 3PL  
The men finished making their yam holes.

Likewise, *jəə* 'wipe'; *jəə + taa* 'erase':

*ε jəə taa na laa kεε-je ke xū baA-peci*  
3SG wipe take away PAST a NOM-say from on piece-book  
He erased a word from the paper.

#### *taa* as relational morpheme<sup>4</sup>

A few verbs require a following object introduced by *taa*: *kəni* 'avoid something'; *mawā* 'avoid a blow'; *mā* 'be discouraged with'; *çəre* 'pass'; *təcā* 'leave someone'. Others are optionally followed by an object introduced by *taa*, and otherwise behave like intransitive verbs: *bata* 'be afraid'; *fiə* 'refuse'; *pəni* 'separate'; *xuru* 'flee'; *māñ* 'precede'. The choice of position for *taa* is fairly free.

*nā kəni laapu na taa lotoo*  
1SG avoid suddenly PAST REL car  
I got right out of the way of the car.

*ri pəni kete taa na rī*  
3PL separate quickly REL PAST 1PL.E  
They quickly separated from us.

<sup>4</sup>*taa* also introduces verbs (as objects) which do not have the nominalising prefix usually found in this context, but can nevertheless be modified by a possessive marker:

*ε fiə taa xwata roo*  
3SG refuse REL listen POSS.REL.2SG  
He refuses to listen to you.

*ε bata taa fe-taa re mwāciri rεε*  
3SG fear REL go-leave POSS.REL country POSS.REL.3SG  
He is afraid to leave his country.

*taa* can also introduce circumstantial modifiers which are not verb arguments:

*are, kə nā kaeε taa re kwereñi taa pe nā*  
tomorrow 2SG DUR gather take away DUR blanket REL bed this  
Tomorrow, you will take the blanket off this bed.

*kə nā fe witaε re mīdəε taa xwāñ*  
2SG DUR go throw DUR trash REL road  
You will go and throw the trash off the road.

Finally, *taa* can be used after stative verbs to express a comparative comparable to the one obtained by the use of *jai* 'pass beyond' (cf. 2.3):

*ε bεññ taa nā = ε bεññ jainā*  
3SG old REL 1SG  
He is older than I.

*nā nɫ təcā re taa rɔ to ɨjɔŋɔ*  
 1SG DUR leave DUR REL 2SG at fork  
 I will leave you at the fork.

Trivalent transitive verbs requiring two objects, one of which denotes the 'dispossessed' participant, will use 'disassociative' *taa* to introduce the latter; these verbs are: *faari* 'ask someone for something'; *pēde* 'steal'; *xanœ* 'ask permission'; *kɔpœ* 'seize'; *pɛ* 'take'; *fatərə* 'demand'; *nāā* 'ask for'; *xādɨ* 'buy'.

*nā nāā çɫɫ-pwī taa rɔ*  
 1SG ask for sucker-banana tree REL 2SG  
 I demand banana-tree suckers from you.

*nā xādɨ ɭaa lotoo taa dape*  
 1SG buy a car REL Dapé  
 I buy a car from Dapé.<sup>5</sup>

### 3. A RELATIONAL MORPHEME OF ADVERBIAL ORIGIN: *ce* 'towards, for'

The relational morpheme *ce* 'towards, for' still behaves like a postverb. It is not possible to say whether it was once a verb, and became a postverb through serialisation, or whether postverbs are a separate source of relational morphemes, distinct from the results of serialisation.

*ce* tends either to form compound verbs<sup>6</sup> by combining with the verb, or contrariwise, to distance itself from the verb and become a relational morpheme.

As a postverb compatible with numerous transitive verbs, *ce* signifies that the action is carried out with a specific purpose:

*nā nārā xiti*  
 1SG think of celebration  
 I think of the celebration.

*nā nārā ce xiti*  
 1SG think of SPECIFIC GOAL celebration  
 I think the celebration over.

*nā taa ce kura*  
 1SG catch SPECIFIC GOAL shrimp  
 I try to catch shrimps.

*ce* allows some intransitive verbs to take an object, and behaves in such cases like a transitivity relational morpheme meaning 'towards, with respect to'. These verbs then express either a movement towards someone or something, or a feeling towards someone: *cə mīā* 'suffer'; *kwāyaa* 'be nostalgic'; *na* 'move towards'; *nārā* 'have pity'; *piçi* 'search'; *tāā* 'stand'; *tɫrɫ* 'be ignorant, fail'.

<sup>5</sup>The same utterance with *xi* 'to' instead of *taa* means: 'I buy a car for Dapé'.

<sup>6</sup>*ce* tends to form compounds with verbs, but can still exceptionally be separated from some by an aspect marker:

*xa* 'speak' *xa-ce* 'call'  
*nā xace na rɔ* (= *nā xa na ce rɔ*)  
 1SG call PAST 2SG  
 I called you.

In the compound *xwi-ce* /do-specific goal/ 'try', however, *ce* cannot be separated from the verb.

*ce* may or may not be included in the verb phrase.

*nã tɫɫɫ ce nĩ-rɔ*  
1SG be ignorant REL name-2SG  
I don't know your name.

*ɲɲãã tãã ce rɔ*  
mummy stand REL 2SG  
Mummy goes forward to meet you.

*ɛ cəmĩã mwãmwaa na ce di re kwãã-re*  
3SG suffer long PAST REL price POSS.REL boat-3SG  
He suffered long (to get) the price of his boat.

#### 4. THE TRANSITIVISING SUFFIX *-ri*

This is the only real verb suffix in Xârâcùù, and is inseparable from the verb it transitivises. It can only be used with a few verbs: *fio* 'refuse', *mãrã* 'worry', *muru* 'live', *kweti* 'be tired', *cara* 'be ashamed', *ɲĩɲãã* 'be complicated'. The object introduced by *-ri* generally denotes the means or the cause:

*nã muru-ri na nũ*  
1SG live-SUF PAST coconut  
I had my fill of coconut.

*nã kweti-ri mãã rɔ*  
1SG be tired-SUF already 2SG  
I am already tired of you.

In all probability, a variant form of this transitive suffix has been incorporated into the verbs *cĩĩĩ* (in free variation with *cĩĩ*) 'distribute, share', and *faari* 'ask for', equivalent to *faa* + the relational morpheme *wã*.

The suffix *-ri* is cognate with several transitive suffixes found in other Caledonian languages (Ajië *-i*, Cèmuhî *-hĩ*), but is no longer productive.

#### 5. THE RELATIONAL MORPHEME *tĩĩ* 'concerning, about'

*tĩĩ* transitivises a few intransitive verbs expressing feelings or acts of communication: *bərə* 'be angry'; *tɛpə* 'make a speech'; *xati* 'quarrel'; *mãã* 'fight'; *tẽĩ* 'cry'.

Its position is variable except after *tẽĩ*: it may be placed immediately after the verb, in which case, if the verb ends in a nasal vowel, the prenasalised form *dĩĩ* will appear; or it may be placed immediately before the object.

*nã mãã dĩĩ na dɔɔ = nã mãã na tĩĩ dɔɔ*  
1SG fight REL PAST earth  
I fought to get land.

*ɛ bərə tĩĩ kwẽẽ-re*  
3SG be angry REL wife-3SG  
He is angry about his wife.



The only possibility with *tē* is to have *dii* immediately after. This suggests compounding has occurred:

*mārādī a tē-dii kāl-re mwiri*  
 snake this cry-REL? skin-3SG ANA  
 Snake cries over his famous skin.

## 6. RELATIONAL MORPHEMES WITH MORE THAN ONE FUNCTION

Relational morphemes can be semantically associated either with arguments or with circumstantial modifiers. In Xârâcùù, some relational morphemes associate with both and have a wide range of functions.

### 6.1 THE RELATIONAL MORPHEME *ŋē*

The relational morpheme *ŋē*<sup>7</sup> can introduce:

- temporal modifiers, in which case it cannot be separated from the object it governs, even when the latter is topicalised; it then means ‘while’:

*ŋē lēēde, ε wā toa = ε wā toa ŋē lēēde*  
 REL evening 3SG ING arrive  
 In the evening, he arrived.

- instrumentals, in which case it is inseparable from its object, and is copied after the verb when the object is topicalised:

*ŋē kwāde, ε wā ça ŋē ri*  
 REL wind 3SG ING hit REL 3PL  
 Using the wind, he starts to beat them.

*ŋē mīi kee-ŋe a, xwi ŋē bwere ŋī-tepə*  
 REL these NOM-say this make REL some bone-speech  
 Using these words, make a few sentences.

- indirect objects of otherwise intransitive verbs: *jana* ‘trade’; *xwiri* ‘sell’; *çōōbo* ‘fool with’.

*ε xwiri ŋē laa nō*  
 3SG sell REL a fish  
 He sells a fish.

In this use, *ŋē* may be placed within the verb phrase or immediately before the object:

*ŋāā jana ŋē na mwe = ŋāā jana na ŋē mwe*  
 mummy trade REL PAST taro  
 Mummy trades her taros.

<sup>7</sup>The range of functions of the relational morpheme *ŋē* is partly comparable to that of the Tongan relational morpheme/postverb/suffix ‘*aki*’ which, according to Churchward (1953), may derive from an earlier verb meaning ‘to use, to wield’. The transitivising suffix ‘*aki*’ can be separated from intransitive verbs by a postverb; after transitive verbs, ‘*aki*’ is an instrumental marker and can be freely placed. See also the various functions of the Paamese case marker *eni*, which has the clitic form *-ni* (Crowley 1983).

When the object is topicalised, the relational morpheme no longer precedes it, but remains in postverbal position; the object is copied by a pronoun after the relational morpheme:

*mĩ nō a, deeri nλ xwiri ŋě re ri*  
 these fish this people DUR sell REL DUR 3PL  
 These fish, people sell them.

– subjects topicalised by postposition: there is thus a confusion in Xârâcùù between the instrumental/object marker and the postposed subject marker, although these functions are usually distinguished in neighbouring languages,<sup>8</sup> where a lexical subject is generally postposed to the verb. In this use, *ŋě* is inseparable from the postposed subject it introduces and is preceded by a slight pause, while the intonation becomes markedly falling.

*a kāmūrū a wā toa*  
 this man this ING arrive  
 This man has arrived.

*ε wā toa ŋě a kāmūrū a*  
 3SG ING arrive S.MKR this man this  
 He has arrived, this man.

The postposed subject may be an independent pronoun:<sup>9</sup>

*nā wā da mōō ŋě gu*  
 1SG ING eat already S.MKR me  
 I have already eaten, myself.

## 6.2 THE RELATIONAL MORPHEME *wā*

Of all the relational morphemes discussed thus far, this one is the most exterior to the verb phrase. It can thus never be separated from the object it introduces. Nevertheless, its close links with the case frame requires that it be dealt with here.

– for some verbs, it introduces the sole argument (usually unmarked): this argument is, as it were ‘peripheralised’, but nevertheless remains in the position usually assigned to the subject. The verbs taking an argument introduced by *wā* express an entirely passive physical state over which the participant has no control:

*taiki çaa*  
 The dog is ill-tempered.

*wā taiki çaa*  
 REL dog bad  
 The dog does not feel well.

The verb *mara* ‘be dizzy, be afraid’ takes a single argument, introduced by *wā*:

<sup>8</sup>Confusion of the two functions is, however, fairly frequent: compare Drehu (Lifou, Loyalty Islands) *hne-*, which marks both the instrumental and the ‘subject’ (agentive) in past constructions.

<sup>9</sup>But the object/possessive form of the pronoun is always used after *ŋě* when it marks an object.

wā pa xūŋfi mara  
REL COLL child be dizzy  
The children are dizzy.

– for other verbs, wā introduces the obligatory object but, unlike other relational morphemes, can never be separated from it:

nā kwāke wā rō  
1SG take care of REL 2SG  
I take care of you.

deeri vire na wā ru (\*deeri vire wā na ru)  
people wrong PAST REL 3DU  
The people wronged them.

– finally, wā may introduce locative or temporal verb modifiers:

wita nō wā kāmīā  
PROHIB stay REL sun  
One should not stay out in the sun.

## 7. VERB COMPOUNDING

Unlike the process of creative relational morphemes, verb compounding in Xârâcùù is centripetal, with verb serialisation giving rise to more or less invariant compounds. There are different degrees of compounding ranging from a construction very close to serialisation to a status close to derivation.

The compound verbs discussed below are all transitive. The first component may be intransitive, in which case the second confers a transitive valency on the whole.

Whatever its original syllable structure, the first component is shortened to CV form. The second is never shortened, but its initial consonant is prenasalised by progressive assimilation whenever the first component has a nasal vowel.

### 7.1 SEMISERIALISATION OR WEAK COMPOUNDING

The compounding process may be weak, insofar as both components may be able to appear separately (in its full, unshortened form, in the case of the first component).

bʌ 'move' (*intrans.*); /ã 'undo'  
bʌ-/ã 'break off'

/ɛ < /ɛɛ 'pull'  
/ɛ-/ã 'unroll (a mat)'

kɔ < kɔu 'scratch'; pɔru 'peel'  
kɔ-pɔru 'peel by scratching'

cu < curu 'tie (a knot)'  
cu-/ã 'untie (a knot)'

Cases of progressive assimilation are:

*mĩ* 'fold'; *-dia* < *tia* 'tear'  
*mĩ-dia* 'tear when folding'

## 7.2 COMPOUNDING FROM A FREE OR SHORTENED FIRST COMPONENT AND A BOUND SECOND COMPONENT

The first component behaves as in 7.1 above, but the second (which is surely of verbal origin) is bound and appears only in compounds. These bound components give rise to a large number of transitive verbs.<sup>10</sup>

*bΛ* 'move' (*intrans.*); *-wi* 'free'  
*bΛ-wi* 'move about to get free'  
*kε-* < *kεε* 'wind' (*intrans.*); *-mwěě* 'close'  
*kε-mwěě* 'cover (said of vines)'  
*mĩ-* < *mĩã* 'urinate' (*intrans.*); *-wi* 'free'  
*mĩ-wi* 'expel in urine'  
*gwe-* < *gwere* 'throw'; *-kəɔ* 'break into pieces'  
*gwe-kəɔ* 'break by throwing'

Progressive assimilation occurs in:

*ʃΛ* 'beat down (sun)' (*intrans.*); *-buru* < *-puru* 'break in two'  
*ʃΛ-buru* 'burn something to break it in two'

## 7.3 COMPOUNDING FROM A BOUND FIRST COMPONENT AND A FREE OR BOUND SECOND COMPONENT

The first component refers to movements made with a part of the body or an instrument. The second is a free or bound transitive verb, and undergoes progressive assimilation.

*mə-* 'with the foot'; *cĩĩ* 'jump'  
*mə-cĩĩ* 'spatter with the foot'  
*kě-* 'with the fingertips'; *-bəru* < *pəru* 'peel'  
*kě-bəru* 'scratch a root with the fingernail to identify it'  
*çə-* 'circular movement of the hand'  
*-puru/-buru* 'break in two'  
*çə-puru* 'break in two by striking with a circular motion of the hand'

<sup>10</sup>The *-wi* suffix meaning 'free' thus transitivises the following verbs: *gə* 'feel nauseous'; *bΛ* 'move'; *pee* 'defecate'; *mĩã* 'urinate':

*bΛ-wi* 'get away from something by squirming'

*bwε bΛ-wi kwifΛi rεε*

crab move-free bond POSS.REL.3SG

The crab squirms to break its bonds.

*-puru* 'break in two'; *-kəɔ* 'break into pieces'; *-mwěě* 'close', etc., behave in the same way.

*kwā-* 'press, pry with a sharp object'; *kə̀ə/-gə̀ə* 'break into pieces'  
*kwā-gə̀ə* 'break with a sharp instrument'

Verb compounding in Xârâcùù is an extremely productive process. Few verbs are unaffected by it, and its flexibility allows an almost unlimited number of new associations to be created. Since all compound verbs are transitive, compounding is not just a means of lexical creation, but also a transitivity process controlled by the second component.

## 8. CONCLUSION

The strength of the bond between the verb phrase and relational morphemes introducing object arguments is variable. The suffix *-ri*, which is inseparable from the verb, and the relational morpheme *wā*, which is inseparable from its object, represent the two extremes, while the remaining relational morphemes have a choice of being included in the verb phrase or preceding their object. In some cases, their verbal origin may explain a preference for inclusion in the verb phrase.

Such inclusion, whether or not it results in transitivity, is a familiar phenomenon in the Oceanic languages. Two hypotheses may be advanced: either these are former suffixes which are reinterpreted as relational morphemes, or on the contrary, relational morphemes have evolved into case markers suffixed to verbs. In Xârâcùù, the two hypotheses are not contradictory. Some relational morphemes clearly originate from verb series (*taa*, *xi*, *fadi*). At the same time, however, the opposite development can be observed with the inclusion of a preposition in the verb phrase when it introduces objects.

There is thus, on the one hand, a bound verb *taa* 'throw, take away', which tends to break off from the verb phrase and turn into a relational morpheme introducing objects (cf. 2.4.2) or even circumstantial modifiers (cf. footnote 4), and on the other, a relational morpheme *ɲɛ* with a variety of functions, which tends to become enclitic to the verb phrase when it introduces objects.

The case of the relational morpheme *tɛ* is less clear: is it a former relational morpheme which is becoming a verb enclitic or rather a former transitive suffix that is becoming a preposition? Comparison of Xârâcùù with the neighbouring languages, Tîrî and Ajië, suggests that the ability to move former transitive suffixes and the creation of an autonomous relational morpheme + object construction outside the verb phrase are innovations in Xârâcùù.

According to Osumi,<sup>11</sup> there are several verb suffixes in Tîrî, which are inseparable from both transitive and intransitive verbs. They have a transitivity effect on intransitive verbs. Semantically speaking, Tîrî verb suffixes correspond more or less to Xârâcùù relational morphemes:

*vai* (dative) 'to, for' corresponds to either *ce* or *tara* in Xârâcùù  
*bɛɛ* 'at, to' corresponds to *xi* or *wā* in Xârâcùù  
*raa* 'off' and *koa* 'apart, away', correspond to *taa* in Xârâcùù.

Tîrî, like Xârâcùù, has several small classes of verbs which are only compatible with one or two suffixes.

In Ajië,<sup>12</sup> there are relatively few relational morphemes or transitivity verb suffixes. The relational morpheme *ye* (goal) corresponds to Xârâcùù *xi* 'to (attributive)':

<sup>11</sup>See Osumi (1990) and Grace (1976).

<sup>12</sup>See La Fontinelle (1976).

*cere daa nãã ye wi' a' nevã ka ə*  
 3PL NEG give REL man this earth QUAL good  
 They did not give good land to this man.

The relational morpheme *ki*, which marks the beneficiary, and the suffix *-i*, which is both a transitivity suffix and an instrumental marker, are part of the verb phrase and appear immediately after the verb. The *-i* suffix corresponds in Xârâcùù to either *-ri* or *taa*, and to *ɲɛ* in its instrumental function:

Ajië:

*na viɔ-i-ə*  
 3SG refuse-SUF-3SG  
 He rejects him, he has had enough of him.

thus corresponds to Xârâcùù:

*ɛ fiɔ-ri ɛ*  
 3SG refuse-SUF 3SG

Ajië:

*cere rhau bara-i wi' a'*  
 3PL all be afraid-SUF man this  
 They are all afraid of this man.

*go ɔɔ-i rhai*  
 1SG flee-SUF lizard  
 I flee the lizard.

are rendered in Xârâcùù by:

*ri xɛɛɔ̃ɔ̃ bata taa kãmurũ a*  
 3PL all be afraid REL man this  
*nã xuru taa ľawa*  
 1SG flee REL lizard

Finally, the Ajië transitivity suffix *-ru* corresponds in Xârâcùù to either *ɔ̃ɔ̃* or the bound element *-wi* in compound verbs (cf. 7.2):

Ajië		Xârâcùù
<i>tãã</i>	'cry' ( <i>intrans.</i> )	<i>tɛ̃</i>
<i>tãã-ru</i>	'cry over'	<i>tɛ̃-dii</i>
<i>gɔ</i>	'vomit' ( <i>intrans.</i> )	<i>gɔ</i>
<i>gɔ-ru</i>	'vomit (something) up'	<i>gɔ-wi</i>

These examples show how hard it can be to find morphosyntactic cognates, even among such closely related languages as Xârâcùù, Ajië and Tiri. These languages have sets of object-introducing relational morphemes and transitivity suffixes which show many semantic and functional similarities, yet are diachronically unrelated, since each language has innovated from its own lexical stock.



Xârâcùù furthermore shows not only the two contrasting tendencies involved in verb serialisation (a centrifugal development creating relational morphemes which separate from the verb phrase, and a centripetal development resulting in the formation of compound verbs), but also another frequent phenomenon in the Oceanic languages, the development of directional verb markers from relational morphemes. While in Tîrî transitivity processes are localised within the verb phrase by inseparable verb suffixes, Xârâcùù seems to be characterised by a cross movement of verbs becoming relational morphemes and relational morphemes yielding verb enclitics.

#### BIBLIOGRAPHY

- CHURCHWARD, C.M.**, 1953, *Tongan grammar*. London: Oxford University Press.
- CROWLEY, Terry**, 1983, Development of a Paamese transitive suffix. In Amran Halim, Lois Carrington and S.A. Wurm, eds *Papers from the Third International Conference on Austronesian Linguistics*, vol. 4: *Thematic variation*, 269-283. PL, C-77.
- GRACE, George W.**, 1975, *Canala dictionary (New Caledonia)*. PL, C-2.  
1976, *Grand Couli dictionary (New Caledonia)*. PL, C-12.
- La FONTINELLE, Jacqueline de**, 1976, *La langue de Houaïlou, Nouvelle-Calédonie: description phonologique et description syntaxique*. Paris: SELAF.
- MOYSE-FAURIE, Claire et Marie-Adèle NÉCHÉRO-JORÉDIÉ**, 1986, *Dictionnaire xârâcùù-français (Nouvelle-Calédonie)*. Nouméa: Edipop.  
n.d., *Grammaire du xârâcùù*. MS.
- OSUMI, Midori**, 1990, *A Grammar of Tinrin (New Caledonia)*. (PhD thesis, ANU, Canberra.)
- RIVIERRE, Jean-Claude**, 1978, Accents, tons et inversion tonale en Nouvelle-Calédonie. *Bulletin de la Société Linguistique de Paris*, 73/1:415-443.

## INCORPORATION OF GENITIVE RELATORS IN THE LANGUAGES OF NEW CALEDONIA AND THE LOYALTY ISLANDS

F. OZANNE-RIVIERRE

### 1. INTRODUCTION

In the languages of New Caledonia and the Loyalty Islands<sup>1</sup> – see map – nouns are divided into two categories: *free nouns*, which can be used on their own, without any affix, and *bound nouns*, which have no lexical autonomy and must be followed by a genitive marker (when quoted out of context) or by a nominal or pronominal possessive determiner.

The free/bound distinction varies from one language to another, and sometimes appears to be arbitrary. We will show, however, that certain factors (lexical replacement, the animate or inanimate nature of the potential possessor) do in fact have a bearing on the assignment of lexical units to one or the other category.

In most languages, the genitive marker which enables bound nouns to be quoted out of context corresponds to the third person singular possessive suffix, but this is not the case in some languages (e.g. Kumak [North] and Drehu [Loyalty]), where the marker has no personal value. The way in which the marker functions also varies from one language to another: sometimes it commutes with all types of possessors (personal pronoun, proper noun or common noun), and sometimes it is retained only in certain contexts (e.g. before common nouns, or before an indeterminate common noun), functioning as a kind of 'relator' (which may indeed have been its original function). In some languages, it may even become incorporated into the root of certain bound nouns, bringing about a restructuring of possessive constructions. We shall see that this reinterpretation of the genitive marker as part of the root mainly affects bound nouns associated with inanimate possessors (parts of a whole, names of parts of plants, anatomical metaphors etc.).

---

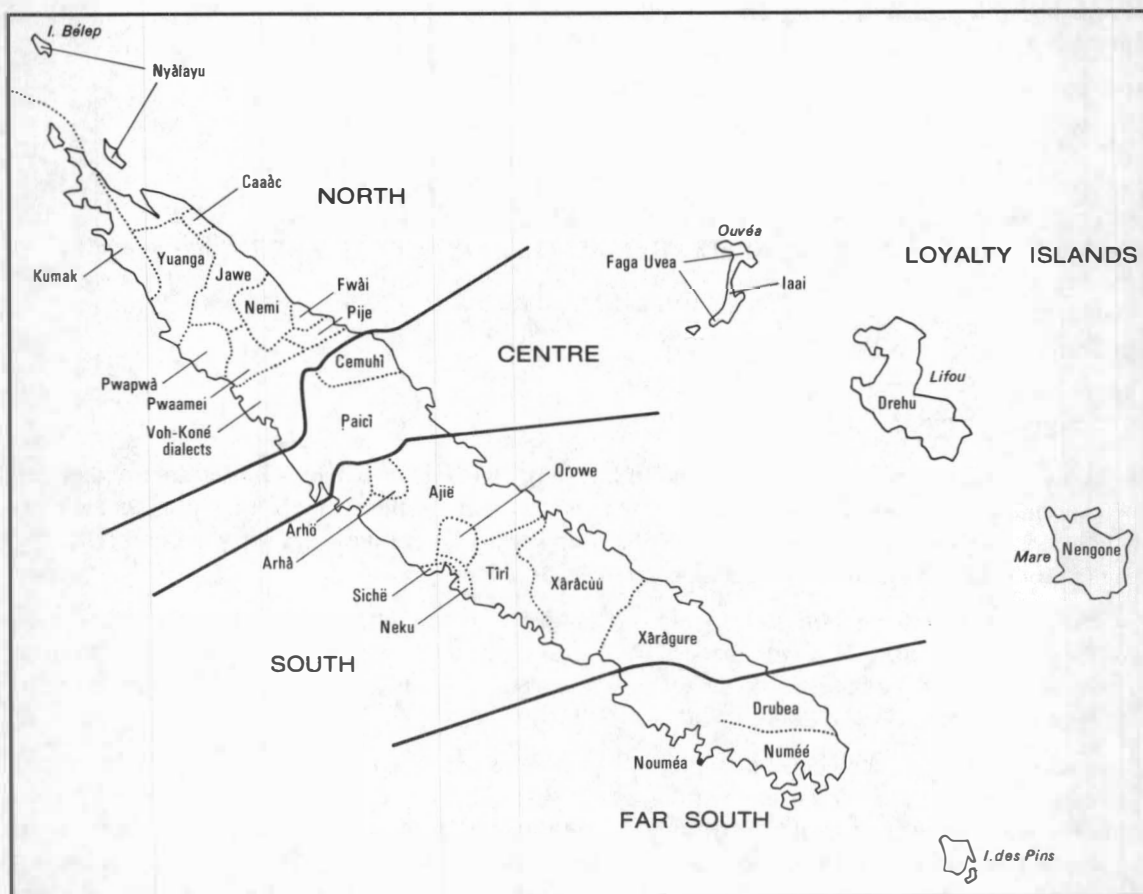
<sup>1</sup>I wish to thank R.A. Blust for his useful comments on an earlier version of this paper. I would also like to thank K.J. Hollyman for giving me access to his article on personalised and non-personalised possession, which also appears in this volume. I have followed his use (in the aforementioned article as well as in previous articles) of the following abbreviations for the names of languages :

Languages of the Loyalty Islands: DEH – Drehu (Lifou), IAI – Iaai (Ouvéa), NEG – Nengone (Maré)

Languages of the mainland: AJE – Ajjé (South), CAC – Caaac (North), CEM – Cèmuhî (Centre), FWA – Fwâi (North), JAW – Jawe (North), KUM – Kumak (North), NMI – Nemi (North), PAC – Paicî (Centre), PAM – Pwaamei (North), PAP – Pwapwâ (North), PIJ – Pije (North), TIR – Tîrî (South), XAC – Xârâcùù (South), YAL – Nyâlâyû (North).

---

Robert Blust, ed. *Currents in Pacific linguistics: papers on Austronesian languages and ethno-linguistics in honour of George W. Grace*, 321-338. *Pacific Linguistics*, C-117, 1991.



THE LANGUAGES OF NEW CALEDONIA

The most commonly used genitive marker in languages which have retained final consonants, the suffix *-n*, may in fact have a double origin: POC *\*-ñá*, third person singular possessive suffix (referring to a definite possessor) and POC *\*ni*, a genitive relator expressing an association between a noun and a non-specific nominal determiner. It will be shown below that the non-personal genitive marker which tends to be incorporated into the root of certain bound nouns is most likely derived from the second of these (the relator *\*ni*). In languages which have lost their final consonants, such as Xârâcüü (South), certain non-etymological nasal vowels may be relics of this relator.

## 2. FREE AND BOUND NOUNS

The majority of bound or 'genitive' nouns belong to the semantic field of nouns traditionally considered as inalienable: kin terms, parts of the human body or of plants, parts of a whole, inherent qualities. But, within this semantic field, we may note that the division between bound and free nouns varies from one language to another, apparently with little semantic motivation. Kin terms or body part terms, bound in one language, may be free in a neighbouring language. Although we cannot disregard the element of arbitrariness which seems to prevail in this field, we can distinguish

two factors which seem to favour the transition from bound to free status: lexical replacement and the animacy of the possessor.

## 2.1 LEXICAL REPLACEMENT

The role of lexical replacement appears clearly in the case of kin terms. In the languages studied here, two sets of kin terms are generally distinguished: names used to refer to a person (bound) and names used to address a person (free). Some languages have retained the clear-cut opposition between the two, and when the term of address is involved in a possessive construction (in familiar speech), the construction is usually indirect with a relator. Here are a few examples taken from two Northern languages (Kumak and Nemi):

(1)	<i>bound nouns</i>		<i>free nouns</i>	
KUM	<i>kibu-t</i>	grandfather (< POC * <i>tumpu-</i> )	<i>huua</i>	grandpa!
	<i>kibu-n</i>	his grandfather	<i>huua i e</i>	his grandpa
NMI	<i>tne-n</i>	his mother (< POC * <i>tina-</i> )	<i>nyaanya</i>	mummy!
			<i>nyaanya ne ek</i>	his mummy
	<i>tnau-n</i>	his father (< POC * <i>tama-</i> )	<i>taara/papa</i>	daddy!
			<i>papa ne e</i>	his daddy

In several languages, certain terms of reference (which generally reflect the Proto Oceanic etymon) have disappeared and been replaced by terms of address. This is what has happened in Cèmuhî, a Centre language, in which several kin terms are now free nouns, indicating possession by means of a relator:

(2)	CEM	<i>nyaa tè-n</i>	his mother	< <i>nyaa</i>	mummy!
		<i>caa tè-n</i>	his father	< <i>caa</i>	daddy!
		<i>ao tè-n</i>	his grandfather	< <i>ao</i>	grandpa!

On the other hand, kin terms which do not have any corresponding terms of address remain bound:

(3)	CEM	<i>iè-n</i>	his sister-in-law	< POC * <i>ipaR</i>
		<i>nai-n</i>	his child	< POC * <i>natu-</i>

Sometimes the terms of reference and of address exist simultaneously:

(4)	CEM	<i>cuö-n</i>	his elder brother	< POC * <i>tuqa-/tuqaka-</i>
	but :	<i>maame</i>	big brother!	and <i>maame tè-n</i> his big brother

The replacement of the reference term by the address term can explain the different treatment in Ajië, a Southern language, of the words 'father' (free noun, indirect possession) and 'mother' (bound noun, direct possession), a fact which Leenhardt (1930:58-62) rather rashly used to back up a hypothesis about the respective importance of 'paternal' and 'uterine' parental links in this society. In fact the *address/reference* distinction only exists for the mother:

(5)	AJE	<i>pani-e</i>	his mother ( <i>bound</i> )	<i>nyaa</i>	mummy!
		<i>pevaaxi-e</i>	his father ( <i>free</i> )	< <i>pevaa</i>	daddy!

Lexical replacement can also explain how a body part term such as 'head', bound in all the Northern languages (6a), has become free in Cèmuhî (6b):

- (6) a. KUM *bwaa-n* his head (*bound*) (*bwaat* top)  
 NMI *bwa-n* his head, its top  
 PAP *gwa-n* his head, its top

but:

- b. CEM *pun* head (*free*) < POC \**pulu* hair  
*puni-n* his head

## 2.2 ANIMATE OR INANIMATE POSSESSORS

The animacy of the potential possessor also influences the transformation of formerly bound nouns into free nouns. The tendency towards lexical independence is indeed more marked when nouns are associated with animate possessors (kin terms, parts of the body) than when they are associated with inanimate ones (parts of a whole, anatomical metaphors).

A good example of this can be found in Drehu (Loyalty). In this language, bound nouns with the genitive marker *-n(e)* are almost all terms which can only have inanimate possessors. Kin terms (reference terms), for example, whose possessors are generally human, have for the most part become free, even when they reflect the Proto Oceanic etymon, whereas in Iaii, another language from the Loyalty Islands, the corresponding terms have remained bound:

- |     |                             |                |                           |
|-----|-----------------------------|----------------|---------------------------|
| (7) | POC                         | IAI            | DEH ( <i>free nouns</i> ) |
|     | * <i>tina</i> -mother       | <i>hinye-n</i> | <i>thin</i>               |
|     | * <i>tama</i> -father       | <i>kame-n</i>  | <i>kem</i>                |
|     | * <i>ipaR</i> sister-in-law | <i>üe-n</i>    | <i>ie</i>                 |

Similarly, in Drehu, body part terms are usually free when their potential possessor is an animate, whereas in metaphorical uses involving an inanimate possessor, they *always* take the genitive marker *-n(e)* which has no personal value. This can be seen from the following examples:

- (8) DEH *ca* foot, *waaca i angeic* his foot (of an animate)  
*ca-n* petiole, stalk (of a plant)  
*he* head, *he i angeic* his head  
*he-* the top (of something)  
*hni* belly, heart (centre of feelings)  
*hni-* the inside (of something)  
*im* arm (< POC \**lima*), *ime i angeic* his arm  
*ime-n* sleeve (of a garment)  
*mek/alaamek* eye (< POC \**mata*)  
*meke-n* point, foremost part  
*qe* mouth, *qe i angeic* his mouth  
*qe-n* orifice, opening

The situation is less clear-cut in Iai, in which some body part terms have remained bound, regardless of whether their possessor is animate or inanimate (cf. 'head' in 9a), while others have become free when applied to an animate possessor and are only bound when applied to an inanimate possessor (cf. 'eye' (< POC \*mata) in 9b and 'foot' in 9c). We may note that in Iai the same genitive marker is used to refer to both animate and inanimate possessors (the possessive suffix *-n*, third person singular):

- (9) IAI a. (*bound noun*) *ba-n* 1. his head (of an animate)  
2. the top (of something)
- b. (*free noun*) *emëk* eye (< POC \*mata)  
*emëka-n* his eye  
(*bound noun*) *hmëka-n* bud (of a plant)
- c. (*free noun*) *ca* foot, leg  
*caa-n* his foot  
(*bound noun*) *caa-n* origin, cause

Similarly, etyma such as POC \*puaq 'fruit' or \*puja 'flower' have two reflexes in Iai and in Drehu: a free form (intransitive verb) and a bound noun form:

- |      |              |                |               |               |
|------|--------------|----------------|---------------|---------------|
| (10) | POC          | IAI            | DEH           |               |
|      | *puaq fruit  | <i>wa</i>      | <i>wa</i>     | to bear fruit |
|      |              | <i>wa-n</i>    | <i>we-n</i>   | fruit         |
|      | *puja flower | <i>hvöng</i>   | <i>eng</i>    | to flower     |
|      |              | <i>vöngö-n</i> | <i>enge-n</i> | flower        |

This lexical divergence from a common root not only involves 'inalienable' terms such as those that refer to human anatomy or to plant parts, it also explains why, in many languages, certain terms which were originally alienable now have two separate reflexes: a free form (for things which can be possessed by a human) and a derived, bound form (with a different meaning) which is associated with inanimates, and which can enter into the formation of compounds. Here are a few examples:

- (11) POC \*Rumaq house
- |         |        | <i>free noun</i> |                   | <i>bound noun</i> |                          |
|---------|--------|------------------|-------------------|-------------------|--------------------------|
|         |        | house            | his house         |                   |                          |
| North   | KUM    | <i>mwa</i>       | <i>mwa-n</i>      | <i>mwa-t</i>      | handle (of a tool)       |
|         | YAL(b) | <i>mwa</i>       | <i>mwa-n</i>      | <i>mwa-r</i>      | handle, container        |
|         | CAC    | <i>mwa</i>       | <i>mwe-n</i>      | <i>mwe-n</i>      | handle, container        |
|         | NMI    | <i>nga</i>       | <i>nge-n</i>      | <i>nge-n</i>      | handle, container        |
|         | PIJ    | <i>nga</i>       | <i>nga-n</i>      | <i>nga-n</i>      | handle, container        |
|         | PAM    | <i>nga</i>       | <i>nga ne-ak</i>  | <i>nga-n</i>      | handle, container        |
| Centre  | CEM    | <i>mwa</i>       | <i>mwa të-n</i>   | <i>mwö-n</i>      | its container            |
|         | PAC    | <i>wá</i>        | <i>wá kê-ê</i>    | <i>wá-rá-</i>     | container-of             |
| South   | AJE    | <i>mwá</i>       | <i>mwá xi-e</i>   | <i>mwá-e</i>      | its container            |
|         | XAC    | <i>mwá</i>       | <i>mwáá-rè</i>    | <i>mwáá-</i>      | container of-            |
|         | TIR    | <i>mwá</i>       | <i>mwá ré-ní</i>  | <i>mwá-ré-</i>    | container-of (something) |
| Loyalty | IAI    | <i>uma</i>       | <i>umwe-n uma</i> | <i>umwe-n</i>     | shelter-for (something)  |



- (12) POC \**saut* spear<sup>2</sup>
- |        |           |                                 |                                   |
|--------|-----------|---------------------------------|-----------------------------------|
|        | spear     | his spear                       | thorn                             |
| KUM    | <i>do</i> | <i>doo-n</i>                    | <i>doo-t</i>                      |
| YAL(b) | <i>do</i> | <i>doo-n</i>                    | <i>doo-r</i>                      |
| CAC    | <i>do</i> | <i>do-n</i>                     | <i>do-n</i>                       |
| NMI    | <i>da</i> | <i>daa-n</i>                    | <i>daa-n</i>                      |
| CEM    | <i>da</i> | <i>da-n</i> (or) <i>da tè-n</i> | <i>da-n</i>                       |
| PAC    | <i>da</i> | <i>da kê-ê</i>                  | <i>daa-ra-</i> (+ inanimate noun) |
| XAC    | <i>jö</i> | <i>jöö-rè</i>                   | <i>jöö-</i> (+ inanimate noun)    |
| TIR    | <i>no</i> | <i>noo-ni</i>                   | <i>noo-</i> (+ inanimate noun)    |
- (13) POC \**salan* road, path
- |     |             |               |               |
|-----|-------------|---------------|---------------|
|     | road        | his road      | track         |
| KUM | <i>daan</i> | <i>dâlâ-n</i> | <i>dâlâ-t</i> |
- (14) POC \**taŋa* basket
- |     |              |                       |                  |
|-----|--------------|-----------------------|------------------|
|     | basket       | his basket            | container        |
| IAI | <i>tang</i>  | <i>tange-n tang</i>   | <i>tange-n</i>   |
| DEH | <i>treng</i> | <i>treng i angeic</i> | <i>treng-e-n</i> |

### 3. PERSONAL POSSESSIVE CONSTRUCTIONS

It is clear from the above examples that direct possessive constructions (*possessed + possessive suffix*) are not necessarily linked with bound nouns.<sup>3</sup> This type of construction is used to indicate possession of the free noun 'house', for example, in most Northern languages (cf. example 11). Notice that, for the same free noun, personal possessive constructions can also vary considerably from one language to another. Here again 'house' is a good example:

- direct construction without inflexion of the final vowel: *mwa*, *mwa-n* (KUM, YAL), *nga*, *nga-n* (PIJ); with inflexion of the final vowel: *mwa*, *mwe-n* (CAC), *nga*, *nge-n* (NMI);
- semi-direct construction with lengthening of the final vowel: *mwâ*, *mwâ-rè* (XAC);
- indirect construction with a relator: *relator + possessive suffix*: *mwa tè-n* (CEM), *relator + object pronoun*: *nga ne-ak* (PAM), *relator + undifferentiated pronoun (object/possessive)*: *wâ kê-ê* (PAC), *mwâ xi-e* (AJE), *mwâ rê-nî* (TIR);
- appositional construction: *classifier + possessive suffix // house*: *umwe-n uma* (IAI, Loyalty).

<sup>2</sup>PMP and POC reconstruction suggested by R.A. Blust (pers. comm.) from the following forms: Puluwa:t *ttow* 'to spear', Palauan *taod* 'three-pronged fish-spear'.

<sup>3</sup>Just as certain free nouns can enter into direct possessive constructions (cf. 'house' in KUM, YAL, CAC, NMI, PIJ, example 11), so can certain bound nouns take indirect constructions with a relator. In TIR, for example, Osumi (1990) distinguishes, alongside 'bound nouns', proper nouns (directly followed by a possessor) and a category of 'linked nouns', which must be linked to a possessor (pronominal or nominal) by the relator *nâ* (e.g. *mwâ-* 'brain', *ao-* 'cheek', *wî-* 'fibres', *wû-* 'shoot', *kû-* 'smoke', etc.).

We should point out that in the other two languages from the Loyalty Islands, DEH and NEG, the personal possessive construction is mixed: direct construction for the first person, indirect construction with a relator for the second and third persons:

(15)		house	my house	your house	his house
	DEH	<i>uma</i>	<i>uma-ng</i>	<i>uma i eö</i>	<i>uma i angeic</i>
	NEG	<i>mma</i>	<i>mma-go</i>	<i>mma ni bo</i>	<i>mma ni bone</i>

As far as the term 'house' is concerned, we may suppose that the direct possessive construction is the oldest.

The appositional construction with a classifier is an innovation characteristic of IAI, which has developed a rich system of possessive classifiers of a typically Micronesian kind (Ozanne-Rivierre 1976:188-195), from a model already existing in all languages of this group, but limited to certain well-defined semantic areas, in particular to food and drink terms (for example, 'his portion of food // yam, taro, banana etc.': 'his drink // water, tea, coffee, etc.').

Indirect constructions with a relator also seem to be more recent in origin:

- In DEH and NEG (Loyalty), the direct construction is only used for the first person (example 15). This form happens to be the only one which has retained a Proto Oceanic reflex (POC \*-*ŋku*). The indirect constructions (*relator + undifferentiated pronoun*) used with the second and third persons are undoubtedly more recent.
- In CEM (Centre), possession of the term 'house' can now only be expressed by means of the relator *tè-* (example 11), direct possession being restricted to the bound derivative *mwö-n* 'its container'. However, the 'spear' *da* (example 12) still bears the trace of the old direct construction *da-n*, alongside the indirect construction *da tè-n*. The modernity of this indirect construction is highlighted by the fact that the same relator *tè-* is used to indicate possession of all recently imported objects (car, radio, guitar etc.).
- Another indication that indirect constructions are more recent is the fact that the relators have a wide variety of forms and therefore must come from different sources. Some of them clearly have a nominal origin (CEM: *tè-* < *dè-n* 'his property', PAC: *kê-* < *kê-é* 'his property'). Others, which involve object pronouns (PAM: *ne-*) or undifferentiated pronouns (AJE: *xi-*, TIR: *rê-*, DEH: *i*, NEG: *ni*) are of more obscure origin (former personal article? former preposition?). It should also be noted that some languages have not just one but several relators available to introduce personal possessors. In some languages the choice of relator is entirely determined by the lexemes whereas in others a certain freedom of choice exists for the same lexeme.

In TIR (South), for example, there are two relators, *rê-* and *nâ-*, whose use is determined as follows (Osuni 1990): the first, *rê-*, has become fixed, and can only be used in combination with certain units ('house', 'boat' and 'cooking pot'). The second, *nâ-*, is productive and is used in combination with an unlimited number of terms.

In CEM (Centre) there is a wider array of relators (*tè-*, *hê-*, *ko-*, *ne-*, *hî-*) and the choice of one or another is sometimes left to the speaker (Rivierre 1980:155). In such cases, naturally, differences in meaning are involved:

(16)	CEM	<i>ani tè-n</i>	his fiancée
		<i>ani hê-n</i>	his desire
		<i>amu tè-n</i>	his country
		<i>amu ne-n</i>	its place (of an object)
		<i>pei tè-n</i>	his stone
		<i>pei hê-n</i>	his sacrum (bone)
		<i>pe-hî-n</i>	its tuber

Moreover, certain relators can also combine with each other. The possessive relator *hê-*, for example, which also functions as a locative preposition ('in') and which indicates a close possessive relationship, can combine with the relator *tè-* (< *dè-n* 'his property'):

(17)	CEM	<i>amo hê-n</i>	its post (of a house)
		<i>amo tè-n</i>	his post
		<i>amo hê-n tè-n</i>	his house post

As we can see, the situation is very complex. In the case of possessive constructions, as in other areas, the languages in this group have each individually introduced a large number of innovations, and they have also borrowed continually from each other. There may also have been external influences, which are hard to evaluate. It would therefore be premature to reconstruct an ancestral possessive system to explain the variety which exists today.

However, I would like to make one further remark, which is important for what is to follow: possessive relators are often clitics to the preceding nouns, and sometimes even tend to merge with them. For example, in the Centre languages (CEM and PAC) which have tones, possessive relators do not have a tone of their own, but are given the same tone as the preceding vowel, although the bound nouns from which they derive have their own tone:

(18)	CEM	<i>dè-n</i> [ <i>dē-n</i> ]	his property	<i>mwa tè-n</i>	[ <i>mwà tè-n</i> ]	his house
				<i>pae tè-n</i>	[ <i>páé té-n</i> ]	his raft
				<i>pei tè-n</i>	[ <i>pēī tē-n</i> ]	his pebble

Similarly, it appears that in XAC (South) the lengthening of the final vowel in the word 'house' before the possessive suffix (cf. example 11, *mwâ*, *mwââ-rè* 'his house') is due to the incorporation of a former relator still attested as *rê-* in the neighbouring language: TIR *mwâ*, *mwâ rê-ni* 'his house'.<sup>4</sup> This incorporation should be considered in parallel with the next question we are going to examine, that of clitic genitive markers in bound nouns.

#### 4. THE THIRD PERSON SINGULAR POSSESSIVE SUFFIX AND THE NON-PERSONAL GENITIVE MARKER

In languages which have retained final consonants (the Northern languages, CEM (Centre) and the Loyalty languages), the POC singular possessive suffixes have reflexes in the form of consonant suffixes. Exceptions to this are DEH and NEG (Loyalty), where only the first person possessive suffix has been retained (cf. example 15).

<sup>4</sup>Lengthening of the vowel before a possessive determiner does not necessarily reflect the incorporation of a relator. It may also reflect an etymological vowel. This can be seen in several languages. Here are a few examples: NMI *do* 'spear', *doo-n* 'his spear' (< POC \**saut*), NMI *do* 'leaf', *doo-n* 'its leaf' (< POC \**ndaun*), XAC *mâda* 'blood', *mâdaa-rè* 'his blood' (< POC \**ndaRaq*), etc.

In languages which have lost their final consonants (PAC (Centre) and the Southern languages), use of consonant suffixes to express personal possession is impossible. These possessive forms have been replaced by syllabic suffixes which generally correspond to object pronouns. Compare KUM, NMI, CEM, IAI, DEH (which have retained final consonants) with PAC, AJE, XAC (which have lost final consonants):

(19)

Possessive	POC	KUM	NMI	CEM	IAI	DEH	PAC	AJE	XAC
1SG	*-ŋku	-ny	-ng	-ng	-k	-ng	-o	-nya	-nâ
2SG	*-mu	-m	-m	-m	-m	(i eö)	-gë	-i	-rö
3SG	*-ña	-n	-n	-n	-n	(i angeic)	-e	-e	-rè

In most languages, the suffix which enables bound nouns to be quoted out of context has the same form as the third person singular possessive pronoun, but in some the genitive marker is different and has no personal value. This is the case in KUM and YAL (North), and in DEH (Loyalty).

#### 4.1 THE NON-PERSONAL GENITIVE MARKER IN KUM AND YAL (NORTH)

Hollyman (in this volume) explains the existence of non-etymological final consonants in the Northern languages as being due to the development of suffixes expressing a very general non-personal possessive relationship, on the pattern of personal possessive suffixation. These non-personal suffixes have various forms and are now merely fossils, entirely bound to their nouns. One of these suffixes, however, has remained productive: the suffix *-t* in KUM, corresponding to *-t/-r* in YAL (*-t* in YAL (a) [Belep] and *-r* in YAL (b) [Balade]).

In these two languages, this suffix has now become the genitive marker for most bound nouns, supplanting the third person singular possessive suffix in this function. However, use of the genitive marker has become more generalised in KUM than in YAL. Kin terms in KUM, for example, are almost always given with this impersonal suffix (*kââma-t* 'father', *kibu-t* 'grandfather', *paabuu-t* 'grandson' etc.), whereas in YAL only a few of these terms accept it (*caama-t* 'father', but *cebo-n* 'grandfather', *paabo-n* 'grandson').

Hollyman also draws attention to the puzzling behaviour of certain non-personal genitive suffixes which, in a given language, are sometimes dropped before a non-specific nominal determiner and sometimes retained and inflected in the same way as an etymological final consonant (KUM *bwaa-t* 'head, top', *bwaa-hoogo* 'summit of a mountain', but *bwara-thi-t* 'nipple'). He also shows that these constructions are not always identical from one language to the other (YAL *bwa-t* 'head, top', *bwala-uuvi* 'top of a yam', but KUM *bwa-kuuvic*).

The KUM dictionary (Haudricourt 1963) shows that the genitive marker *-t* always commutes with personal possessive suffixes, whereas in front of a nominal determiner it can behave in one of three ways:

- It can be dropped, and this is the most common construction (*yada-t* 'thing', *yada-n* 'his personal belongings' and *yada-wany* 'the equipment of a boat'; *duaxa-t* 'back', *duaxa-n* 'his back' and *duaxa-mwa* 'the wall of a house'; *dexa-t* 'juice, liquid' and *dexa-ciic* 'sap' ('juice-tree').

- It can be retained (*kuuaa-t* 'foot', *kuuaa-n* 'his foot' and *kuuaa-t taap* 'table leg'; *pap*, *pavu-t* 'nest', *pavu-la* 'their nests' and *pavu-t mââlic* 'bird's nest' (but *pa-mabo* 'wasp's nest!'); *mwa-t* 'handle' (of an implement) and *mwa-t gi* 'axe handle'.
- More rarely, it can inflect to *r-* (*bwaa-t* 'top', *bwaa-n* 'his head', *bwara-thi-t* 'nipple' (but *bwaa-hoogo* 'summit of a mountain'); *wa-t* 'vein, cord' (<POC \**uRat*), *wara-nep* 'reef of a sail' (cord of-sail).

Only more detailed descriptions and deeper comparative investigation will help us to understand the reasons for these variations. In the meantime we propose simply to compare the non-personal genitive suffix (-*t* in KUM, -*t/-r* in YAL) with the impersonal possessive relator (*rV*) attested in PAC (Centre). From this comparison it will become obvious that the non-personal genitive marker in KUM and YAL, which has sometimes been incorporated into the root, is probably a reflex of a former impersonal possessive relator. In PAC, impersonal possessors are introduced by means of a relator with the form *rV*, *V* being a vowel of the same timbre as the final vowel of the thing possessed (Bensa and Rivierre 1976; Rivierre 1983).

Among personal possessors we find: possessive suffixes, proper nouns and common nouns preceded by a personal article (*pwi* 'singular', *du* 'feminine dual' and *tupè-du* 'masculine dual', *pa* 'feminine plural' and *tee-pa* 'masculine plural').

Among impersonal possessors we find: common nouns without an article (having a generic meaning and entering into the formation of compounds) or preceded by impersonal articles (*i* 'definite', *cè* 'indefinite'). Here are some examples of possessive constructions with bound nouns:

(20)	PAC	<i>personal possessor</i>		<i>impersonal possessor</i>	
		<i>î-é</i>	his arm/hand	<i>î-rî göö</i>	crab's claw
		<i>au-é</i>	his track	<i>au-ru wâ</i>	house site
		<i>â-é</i>	his leg/foot	<i>â-râ wâ</i>	house post
		<i>cèù-é</i>	his back	<i>cèù-rù wâ</i>	back of a house
		<i>cèù Têâ</i>	Têâ's back	<i>cèù-rù i aboro</i>	man's back
		<i>cèù pwi aboro</i>	man's back		( <i>i</i> = impersonal article)
			( <i>pwi</i> = personal article)		

Some free nouns are constructed on the pattern of bound nouns:

(21)	PAC	<i>né</i>	name	<i>néé-ré wââo</i>	lineage name
		<i>néé-gë</i>	your name	<i>néé-ré i aboro</i>	the man's name

But more often than not they are introduced by means of the relator *kä-/kë-/kê-* (+ personal possessor), *kärä* (+ impersonal possessor):

(22)	PAC	<i>wâ</i>	house
		<i>wâ kê-é</i>	his house
		<i>wâ kärä i aboro</i>	the man's house

We have already seen in examples (11) and (12) that the POC etyma \**Rumaq* 'house' and \**saut* 'spear' have two reflexes in most New Caledonian languages: a free form, respectively 'house' and 'spear', and a bound derivative, 'container' and 'thorn', whose possessor is usually inanimate, and which enters into the formation of numerous compounds. It is likely that the non-personal genitive marker in KUM (-*t*) and in YAL (-*t/-r*), which characterises these bound derivatives and tends to be



incorporated into them, is related to the PAC impersonal possessive relator *rV*. Consider the following PAC and YAL (b) examples:

(23) *bound derivatives*

	KUM	YAL (b)	PAC	
house	<i>mwa</i>	<i>mwa</i>	<i>wâ-</i>	house
container		<i>mwa-r</i>	<i>wâ-râ-</i>	container-of-
bottle		<i>mwâlâ-we</i>	<i>wâ-râ-wado</i>	bowl
spear	<i>do</i>	<i>do</i>	<i>da</i>	
thorn	<i>doo-t</i>	<i>doo-r</i>	<i>da-ra-</i>	thorn-of-

The same parallelism appears for bound nouns with inanimate possessors, but, as we pointed out earlier on, with differences in construction between KUM and YAL:

(24) *bound nouns*

	KUM	YAL (b)	PAC	
fruit	<i>pwâ-t</i>	<i>pwa-r</i>	<i>pwâ-râ-</i>	fruit-of-
coconut	<i>pwâ-nu</i>	<i>pwâ-nu</i>	<i>pwâ-râ-nû</i>	
point	<i>mââ-t</i>	<i>maa-r</i>	<i>mâ-râ-</i>	point-of-
spear point		<i>mââlâ-do</i>	<i>mâ-râ-pwadù</i>	banana shoot
inside	<i>na-t</i>	<i>na-r</i>	<i>nâ-râ</i>	inside-of-
hold (of a boat)	<i>na-waany</i>	<i>nâlâ-waang</i>	<i>nâ-râ-wâ</i>	room (inside-of-house)

There clearly seems to be some kinship between the non-personal possessive suffix from the Far North (KUM and YAL) and the impersonal genitive relator in PAC (Centre). The main function of this genitive marker may originally have been to associate a noun with a non-specific nominal determiner. It is widely used in PAC to form compound nouns (Noun + *rV* + Noun without article) and may, within this language, have spread to all nominal possessors, animate and inanimate, definite or indefinite, unpreceded by personal articles. In KUM and YAL it has become a neutral genitive marker, enabling most bound nouns to be quoted, but it has become especially widespread in KUM, affecting even kin terms (cf. above).

The origin of this genitive marker is unclear. It may derive from the locative preposition common to all Northern languages (CAC *re*, PAM, NMI, PIJ, JAW, FWA *le* 'in'), which is used to introduce only inanimate nouns without articles (CAC *re we*, NMI *le we* 'in the water'). It is also sometimes used in these languages as a genitive relator before a non-specific determiner: CAC *we re kuc* 'sugarcane juice' ('water/ in/ sugarcane'), NMI *cee le gi* 'axe handle' ('wood/ in/ axe'). But whatever the source of this non-personal genitive marker may be, we will see that its functioning very closely echoes that of the impersonal possessive suffix *-n(e)* found in DEH (Loyalty).

## 4.2 THE IMPERSONAL GENITIVE MARKER IN DEH (LOYALTY ISLANDS)

We saw in section 2.2 that the class of bound nouns in DEH consists mostly of nouns associated with inanimate possessors (cf. examples 7, 8 and 9) and that the suffix *-n*, which enables these bound nouns to be quoted, has no personal value (cf. example 8). In fact, bound and free nouns are



distinguished above all by the presence or absence of this suffix in quoted forms, while their possessive constructions are very similar (Moyses-Faurie 1983:59-61).

With bound nouns, the impersonal genitive suffix *-n* is dropped before personal possessors (pronouns and proper nouns). Possessors in the first person (singular, dual and plural) have direct suffixation. Possessors in the second and third persons and proper nouns are introduced by the relator *i*. Before common noun possessors, the genitive suffix is retained, and an eliding vowel is developed (*-ne*). Its functioning is very similar to that of the 'construct suffix' (Lichtenberk 1985:100) attested in several Micronesian languages:

(25)	DEH	<i>pengö-n</i> (bound noun)	the way one is (or does things)
		<i>pengö-ng</i>	the way I do things
		<i>pengö i eö</i>	the way you do things
		<i>pengö i angeic</i>	the way he/she does things
		<i>pengö i Wamo</i>	the way Wamo does things
		<i>pengö-ne la nöj</i>	the custom of this country

With free nouns, the personal possessive construction is identical to that used for bound nouns. As for common noun possessors, they are introduced by means of the relator *ne*:

(26)	DEH	<i>uma</i> (free noun)	house
		<i>uma-ng</i>	my house
		<i>uma i eö</i>	your house
		<i>uma i angeic</i>	his/her house
		<i>uma i Wamo</i>	Wamo's house
		<i>uma ne la qatreföe</i>	the old woman's house

The relator *ne* is also used to introduce non-specific nominal determiners and to mark destination, thus clearly reflecting the POC associative relator *\*ni* (Hooper 1985, Lichtenberk 1985):

(27)	DEH	<i>waaca ne gutu</i>	chicken leg (leg/of/chicken)
		<i>uma ne drai</i>	meeting house (house/for/meeting)

The relator *ne* used to introduce definite possessors of free nouns in DEH (example 26: *uma ne la qatreföe* 'the old woman's house') is also a reflex of POC *\*ni* with a wider syntactic function. But what about the impersonal genitive suffix *-n(e)* of bound nouns (example 25)? Is it a reflex of the POC third person singular possessive suffix *\*-ña* which has lost all personal meaning, or of the impersonal genitive relator *\*ni*? The answer to this question is not obvious, but, in so far as this suffix mainly applies to bound nouns associated with inanimates (cf. example 8), where it has no personal value and where it is retained before non-specific nominal determinants, it would seem more likely to be a reflex of the associative relator, rather than being a relic of some personal possessive suffix or other:

(28)	DEH	<i>meken</i>	point, foremost part (< POC <i>*mata</i> )
		<i>mekene-hee</i>	prow of a boat
		<i>wen</i>	fruit, grain, drop (< POC <i>*puaq</i> )
		<i>wene-sinoë</i>	fruit in general (fruit/tree)
		<i>wene-leemen</i>	lemon
		<i>wene-ngöni</i>	grain of sand
		<i>wene-mani</i>	raindrop

<i>qen</i>	opening, orifice ( <i>qe/qĕ</i> mouth)
<i>qene-hnaop</i>	hole
<i>qene-hnafij</i>	nostril
<i>qene-taaxoj</i>	anus
<i>trengen</i>	container ( <i>treng</i> basket < POC * <i>taŋa</i> )
<i>trengene-tim</i>	carafe (container/water)
<i>trengene-an</i>	stomach (container/food)

In the last example, *trengen* 'container', the incorporation of the suffix is so complete that it is possible, from this derived form, to reconstruct the following personalised possessive construction:

- (29) DEH        *trengene i angeic*                      its container (for the shell of a shellfish)  
                   as opposed to: *treng e i angeic*                      his basket (of a human being)

This reinterpretation of the genitive marker as forming an integral part of the word is strongly reminiscent of the situation we described above in KUM. Compare the DEH example (29) with the following example from KUM:

- (30) KUM        *mâât (mââ-t)*                                      point (of something) < POC \**mata*  
                   *mââr i nya*    my point (e.g. the point of my spear)

The impersonal genitive markers found in Drehu and Kumak are reminiscent of the suffixes *-i* and *-gi* described by Codrington (1885) in languages from the north of the New Hebrides. These suffixes, which characterise the 'free' forms of inalienable nouns (e.g. Mota: *mata-i* 'eye' (in general) as opposed to *na mata-na* 'his eye') are, according to Hooper (1985), reflexes of a former POC relator \**qi*, used to introduce non-specific determiners of inalienable nouns.

In Motlav, there are two suffixes with arbitrary distribution which enable inalienable nouns to be used independently: *-ge* (<POC \**qi*) and *-n*, whose origin is debatable (POC personal possessive suffix \**-ñā* or POC associative relator \**ni*). The same uncertainty was noted earlier on for the impersonal genitive suffix *-n(e)* in Drehu, and in this case we expressed our preference for a non-personal source (POC \**ni*). To come back to Motlav, Codrington (1885:141) underlines the fact that speakers perceive the suffixes *-ge* and *-n* of free forms (*tqe-ge* 'belly', *tri-ge* 'body', *nte-n* 'child', *tle-n* 'egg') as both having the same generalising effect, and the suffix *-n* in such contexts as being quite distinct from the third person singular possessive suffix.

Let us now see what the situation is in those mainland languages which have retained final consonants, and in which the standard genitive marker for bound nouns (*-n*) corresponds to the third person singular personal possessive suffix.

## 5. SPECIFIED AND UNSPECIFIED POSSESSION IN NMI (NORTH)

In the languages of the Northern part of mainland Caledonia, with the exception of KUM and YAL, the genitive marker which enables bound nouns to be quoted out of context corresponds to the third person singular possessive suffix *-n*. This is the case in Nemi, where the genitive suffix is dropped when followed by a personal possessor (pronouns and proper nouns) and by a definite nominal possessor (animate or inanimate):

(31)	NMI	<i>cii-n</i> <i>cii Kaavo</i> <i>cii vi hnook</i> <i>cii vi ceek</i>	his skin, its bark Kaavo's skin the woman's skin the bark of the tree
------	-----	---	--

With an indeterminate possessor, two types of genitive construction occur (Ozanne-Rivierre 1979, 1:44-45):

1. – Restoration of an old final consonant (<POC \*-C-) + *e* + a non-specific determiner:

(32)	NMI	<i>tnamaa-n</i> <i>tnamaa vi ceek</i> but : <i>tnamaye-ceek</i>	his eye, its bud < POC * <i>mata</i> eye the bud of the tree tree bud ( <i>unspecified</i> )
(33)	NMI	<i>nai-n</i> <i>nai ven daama</i> but : <i>naye-daama</i>	his child < POC * <i>natu</i> the chief's child a chief's child ( <i>unspecified</i> )
(34)	NMI	<i>puu-n</i> <i>puu vi ceek</i> but : <i>puxe-ceek</i>	its origin, its base < POC * <i>puqun</i> the base of the tree tree stump ( <i>unspecified</i> )

2. – Preservation of the suffix *-n* + non-specific determiner:

(35)	NMI	<i>cii-n</i> <i>cii vi ceek</i> but : <i>cii-n ceek</i> <i>cii-n hwa-n</i> <i>cii-n jigo</i>	his skin, its bark < POC * <i>kulit</i> the bark of the tree bark (in general) ( <i>unspecified</i> ) his lips (skin-of/mouth-his) 1. mangrove bark 2. name of a fish ( <i>Mugil</i> sp.)
(36)	NMI	<i>hi-n</i> <i>hi vi ceek</i> but : <i>hi-n ceek</i> <i>hi-n pwec</i>	his arm, his hand the branch of the tree branch of a tree ( <i>unspecified</i> ) tributary of a river
(37)	NMI	<i>pwe-n</i> but : <i>pwe-n ceek</i> <i>pwe-n pijing</i> <i>pwe-n kut</i>	its fruit (of a specific tree) < POC * <i>puaq</i> fruit (in general) ( <i>unspecified</i> ) banana raindrop
(38)	NMI	<i>hwa-n</i> <i>hwa vi hnook</i> but : <i>hwa-n nga</i> <i>hwa-n daaoot</i> <i>hwa-n dawe</i>	his mouth, its orifice, its hole < POC * <i>papaq</i> the woman's mouth door (hole-of/house) ( <i>unspecified</i> ) estuary (hole-of/river) clearing (hole-for/breeze)

The suffix *-n* which occurs in non-specific genitive constructions apparently has nothing to do with the personal possessive suffix *-n* (< POC *n̄a*), but would rather seem to be a reflex of the POC associative relator \**ni*. The suffix *-n* which appears in bound derivatives associated with inanimate possessors (examples 11 and 12) probably has the same impersonal origin:

- |      |       |   |  |
|------|-------|---|--|
| (39) | NMI   | <i>nga</i><br><i>nge-n</i><br><i>nga vi hnook</i>   | house < POC * <i>Rumaq</i><br>his house<br>the woman's house   |
|      | but : | <i>nge-n + indeterminate noun</i><br><i>nge-n suma</i><br><i>nge-n yaak</i><br><i>nge-n jixet</i>               | container-of-<br>lungs (container-of/spit)<br>match (container-of/fire)<br>bow (container-of/arrow)  |
| (40) | NMI   | <i>do</i><br><i>doo-n</i><br><i>doo vi kac</i>  | spear < POC * <i>saut</i><br>his spear<br>the man's spear  |
|      | but : | <i>doo-n dogan</i>  | orange-tree thorn  |
| (41) | NMI   | <i>ceek</i><br><i>cee-n + indeterminate noun</i><br><i>cee-n nga</i><br><i>cee-n tnek</i><br><i>cee-n cilek</i> | tree < POC * <i>kai</i><br>wood-of-, wood-for-<br>(roof) beams (wood-for/house)<br>firewood (wood-for/oven)<br>cooking wood (wood-for/cooking pot) |
| (42) | NMI   | <i>wâk</i><br><i>wâxe-n</i>   | vein, cord, creeper < POC * <i>uRat</i> nerve, vein<br>my vein   |
|      | but : | <i>wâ-n + indeterminate noun</i><br><i>wâ-n jic</i><br><i>wâ-n pwe</i>  | cord-of-<br>belt (cord-of/belly)<br>fishing line (line-of/fishing rod)   |

## 6. NON-ETYMOLOGICAL NASAL VOWELS IN XAC (SOUTH)

We have already seen that, in most New Caledonian languages, possessive relators are often clitics to the preceding word. When a genitive marker is incorporated into the root of certain lexical units, we prefer to consider it as the reflex of a genitive relator rather than of a personal possessive suffix. It so happens that this incorporation mainly affects bound nouns associated with inanimate 'possessors' (parts of a whole, anatomical metaphors). Now, there is one point on which all descriptions agree: bound nouns referring to inanimate possessors may accept the addition of a third person singular possessive suffix, but only in specific enunciative contexts, in which the inanimate possessive determiner is known, visible, or has previously been mentioned. In fact, these bound nouns, when quoted out of context, do not usually take a personal possessive suffix, but are rather set in a nominal group similar to a compound, in which the noun is associated with its most general determiner.

In NMI (North), for example, the general term for 'fruit', 'branch', 'bark' is not the personal possessive form *pwe-n*, *hi-n*, *cii-n*, but a form implying association with the most general determiner, that is, *ceek* 'tree': *pwe-n ceek* 'fruit', *hi-n ceek* 'branch', *cii-n ceek* 'bark' (cf. examples 35, 36 and 37).

In XAC (South), Moysse-Faurie (n.d.) notes that the third person singular possessive suffix *-rè* enables bound nouns in general to be quoted. But the most neutral way of using bound nouns associated with inanimates independently is by affixing a nominal determinant with a very general sense, such as *döu* 'thing' or *kètè* 'place'. A bound noun such as *ké-* 'handle', for instance, will be quoted as *ké-döu* (handle-thing) if one wants to refer to a handle in general, and as *ké-rè* (handle-its)

to refer to the handle of a particular tool. This unwillingness to use a personal possessive suffix to cite bound nouns associated with inanimates out of context reinforces my hypothesis that the genitive marker, sometimes incorporated into this type of noun, is more likely to be a neutral genitive marker (associative relator) than a personal possessive suffix.

If certain non-etymological final consonants in the Northern languages derive from the incorporation of such relators, then it is possible that some non-etymological nasal vowels found in Southern languages (where final consonants have been lost) are also a relic of the associative relator, which appears in several Northern languages as the suffix *-n* (cf. section 5), and in DEH (if my hypothesis is correct) as the impersonal genitive marker *-n(e)* (cf. section 4.2).

Loss of final consonants is one of the main sources of phonological nasal vowels, by anticipatory assimilation, in the Southern languages:

(43)	POC		NMI (North)	XAC (South)
	<i>*poñu</i>	turtle	<i>pwen</i>	<i>pwé</i>
	<i>*quma</i>	to plough	<i>hum</i>	<i>xû</i>
	<i>*lajo</i>	fly	<i>nen</i>	<i>nâ</i>
	<i>*waŋka</i>	boat	<i>wang</i>	<i>kwâ</i>

But certain nasal vowels attested in bound nouns are not etymological, and these undoubtedly reflect the incorporation of a genitive marker which, in our opinion, is unlikely to come from a personal source, since the bound nouns in question are more often than not associated with inanimates:

(44)	XAC	<i>pwâ-</i> <i>pwâ-kwââ</i> <i>pwâ-mêdè</i> <i>pwâ-kwé</i>	fruit of, drop of < POC <i>*puaq</i> fruit (in general) (fruit-tree) orange (fruit- <i>Citrus</i> ) drop of water (drop-water)
(45)	XAC	<i>nyî-</i> <i>nyî-ji</i> <i>nyî-nû</i> <i>nyî-kùrè</i>	juice of < POC <i>*suRuq</i> mother's milk (juice-breast) coconut milk (juice-coconut) sauce (juice-cooking pot)
(46)	XAC	<i>nê-</i> <i>nê-kwââ</i> <i>nê-nû</i>	leaf of < POC <i>*ndaun</i> leaf (in general) (leaf-tree) coconut palm (leaf-coconut tree)
(47)	XAC	<i>kê-</i> <i>kê-döu</i> <i>kê-giè</i> <i>kê-puè</i> <i>kê-nè</i>	handle of, wood of < POC <i>*kai</i> tree, wood handle (in general) (handle-thing) axe handle (wood-axe) stake to hold a net (wood-net) match (wood-fire)
(48)	XAC	<i>kwâ-</i> <i>kwâ-kwâ</i> <i>kwâ-chii</i>	cord of, bond of < POC <i>*uRat</i> nerve, vein rigging of a boat (cord-boat) fishing line (cord-to fish)

(49)	XAC	<i>xwâ-</i>	hole of, opening of < POC * <i>papaq</i>
		<i>xwâ-kû</i>	nostrils (hole-nose)
		<i>xwâ-döu</i>	wound (hole-thing)
		<i>xwâ-kwè</i>	spring (hole-water)

Comparison with the non-specific possessive constructions illustrated previously (some of them with the same etyma) in NMI (North), examples (35) to (42), and in DEH (Loyalty), example (28), would seem to provide convincing evidence that the non-etymological nasal vowels found in XAC are again relics of the genitive relator POC \**ni*.

## 7. CONCLUSION

Hollyman (in this volume), taking up the problem posed by Grace, in 1972, of canonical shapes in the languages of Northern New Caledonia, shows that certain non-etymological final consonants may derive from suffixes indicating non-personal, non-individualised possession. This hypothesis has been entirely confirmed here in the case of the suffix *-t* in KUM (*-t/-r* in YAL), the source of which is probably a genitive relator (cf. section 4.1). It is harder to trace the origin of the other non-etymological final consonants which have very varied forms and are more fixed. More detailed comparative analyses will perhaps enable us one day to find out where these 'intruders' came from. However, the NMI example in (42) is interesting, because the suffix *-k* in *wâ-k*, which corresponds to *wa-t* in KUM (< POC \**uRat* 'nerve, vein') may derive from the relator *ko/xo* 'on', which sometimes functions as a possessive relator (NMI *daama* 'chief', *daama xo-ng* 'my chief', CEM *cinu ko-n* 'his illness' (illness on-him), or as an associative relator (NMI *wâ-k* 'vein, cord, creeper', *wâ xo hiu-ng* 'the vein of my hand' (vein/on/hand-my)). But this source, plausible for the suffix *-k* in *wâ-k*, is much more doubtful in the case of the non-etymological consonant *-k* found in forms such as *ceek* 'tree' (< POC \**kai*), *ciik* 'louse' (< POC \**kutu*), *kuuk* 'yam' (< POC \**qupi*). There may be several explanations for these anomalies in the canonical forms. It may be the case that final consonants attested in the Caledonian North derive at the same time from the preservation of certain POC final consonants, from the incorporation of impersonal genitive markers (cf. Hollyman), and also from the spontaneous closing of syllables in certain contexts (after closed vowels).<sup>5</sup> In any case, we have seen traces of the incorporation of genitive relators, which has probably been a recurrent process in this group of languages as a whole, not only in languages with final consonants, but also in the non-etymological nasal vowels of the Southern languages.

## BIBLIOGRAPHY

- BENSA, A. and J.C. RIVIERRE, 1976, De quelques genres littéraires dans la tradition orale Paicî (Nouvelle-Calédonie), *Journal de la Société des Océanistes* 32/50:31-66.
- BLUST, R.A., 1978, Review of T.A. Sebeok, ed. *Current trends in linguistics*, vol.8: *Linguistics in Oceania*. The Hague, 1971. *Language* 54/2:467-480.
- BURLING, R., 1966, The addition of final stops in the history of Maru (Tibeto-Burman). *Language* 42/3:581-587.
- CODRINGTON, R.H., 1885, *The Melanesian languages*. Oxford: Clarendon Press.
- DUBOIS, M.-J., n.d., Dictionnaire nyâlayu-français (Belep), Nouvelle-Calédonie. MS.

<sup>5</sup>Epentheses of this kind after high vowels (*-k* after *u*, *-t* after *-i*) have been put forward by Burling (1966) to explain aberrant syllabic correspondences in Maru, a Tibeto-Burman language. Blust (1978:474-475) suspected a similar origin for certain non-etymological final consonants in New Caledonian languages.



- GRACE, G.W., 1972, On canonical shapes in the language of Koumac. In J.M.C. Thomas and L. Bernot, eds *Langues et techniques, nature et société (Hommage à A.G. Haudricourt)*, 135-140, vol.1. Paris: Klincksieck.
- 1975, *Canala dictionary (New Caledonia)*. PL, C-2.
- 1976, *Grand Couli dictionary (New Caledonia)*. PL, C-12.
- HAUDRICOURT, A.-G., 1963, *La langue des Nenemas et des Nigoumak (dialectes de Poum et de Koumac, Nouvelle-Calédonie)*. Te Reo Monographs. Auckland: Linguistic Society of New Zealand.
- HAUDRICOURT, A.-G. and F. OZANNE-RIVIERRE, 1982, *Dictionnaire thématique des langues de la région de Hienghène : pi je, fwâi, nemi, jawe*. Paris: SELAF.
- HOLLYMAN, K.J., n.d., *Dictionnaire caaàc-français*. MS.
- 1991, Personalised and non-personalised possession: final consonants in Kumak and other languages of Far Northern New Caledonia.
- HOOPER, R., 1985, Proto-Oceanic \*qi. In Pawley and Carrington, eds 1985:141-167.
- LA FONTINELLE, J. de, 1976, *La langue de Houailou*. Paris: SELAF.
- LEENHARDT, M., 1930, *Notes d'ethnologie néo-calédonienne*. Paris: Institut d'Ethnologie.
- 1946, *Langues et dialectes de l'Austro-Mélanésie*. Paris: Institut d'Ethnologie.
- LICHTENBERK, F., 1985, Possessive constructions in Oceanic languages and in Proto-Oceanic. In Pawley and Carrington, eds 1985:93-140.
- MOYSE-FAURIE, C., 1983, *Le drehu, langue de Lifou (Iles Loyauté)*. Paris: SELAF.
- n.d., *Grammaire du xârâcùù (Nouvelle-Calédonie)*. MS.
- MOYSE-FAURIE, C. and M.A. NÉCHÉRO-JORÉDIÉ, 1986, *Dictionnaire xârâcùù-français (Nouvelle-Calédonie)*. Nouméa, Edipop.
- OSUMI, Midori, 1990, A grammar of Tinrin, New Caledonia. (PhD thesis, ANU, Canberra).
- OZANNE-RIVIERRE, F., 1976, *Le iaai, langue mélanésienne d'Ouvéa, Nouvelle-Calédonie: phonologie, morphologie, esquisse syntaxique*. Paris: SELAF.
- 1979, *Textes nemi*. 2 vols. Paris: SELAF.
- OZANNE-RIVIERRE F. and M. MAZAUDON, 1986, *Lexique nyâlayu-français (Balade), Nouvelle-Calédonie*. Paris: LACITO-CNRS.
- PAWLEY Andrew and Lois CARRINGTON, eds, 1985, *Austronesian linguistics at the 15th Pacific Science Congress*. PL, C-88.
- RIVIERRE, J.C., 1980, *La langue de Touho: phonologie et grammaire du cémuhi (Nouvelle-Calédonie)*. Paris: SELAF.
- 1983, *Dictionnaire paici-français*. Paris: SELAF.
- n.d., *Dictionnaire cèmuhi-français*. MS.
- SAM, L.D., 1980, *Lexique Lifou-français*. Nouméa: CTRDP.

## HOW TO TALK CRICKET: ON LINGUISTIC COMPETENCE IN A SUBJECT MATTER

ANDREW PAWLEY

England soon lost Fowler. But...Robinson made certain of a lengthy lead on first innings with a determined maiden Test hundred. In eight and three-quarter hours at the crease, with seventeen 4s, he gave his only chance when he was 54, Kirmani missing stumping him off Sivaramakrishnan's googly, which almost hit the off stump. Eliminating the threat of the occasional low bounce by wary play against short balls on the stumps, Robinson shared three-figure stands with Lamb and Downton before, early on the fourth day, he was caught at the wicket off his gloves from a ball from Kapil Dev that lifted unexpectedly off a length.

Most of Robinson's runs came off his legs and behind square on the off side, where by opening the blade he repeatedly found gaps; but when the spinners over-pitched, he off-drove with certainty and power....Sivaramakrishnan polished off the innings...completing his third successive bag of six wickets in the series.

(*Wisden cricketers' almanack* 1986, pp. 888-889, reporting a match between England and India)

### 1. INTRODUCTION<sup>1</sup>

Here I take up a question that has engaged George Grace in two of his books, *An essay on language* (1981) and *The linguistic construction of reality* (1987): what else is there to a language besides the kinds of knowledge represented in grammars and dictionaries?

Building on ideas adumbrated by Grace, I will argue that to speak and understand a language competently we need to command (among other things) a variety of *subject matter codes*. I use 'subject matter code' to refer to the conventions shared by members of a speech community that specify, in more or less detail, what things may be said about a particular subject or topic, how these things are said, idiomatically, and when and why they are said, appropriately. That is to say, it is a code for binding linguistic content with form, context and purpose. An example of discourse in a

---

<sup>1</sup>I am indebted to Scott Allan, Robert Blust, Koenraad Kuiper, Frank Lichtenberk, John Lynch, Jim Martin, Lawrence Reid and Frances Syder for critical comments and suggestions on a draft of this paper.

Robert Blust, ed. *Currents in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, 339-368. *Pacific Linguistics*, C-117, 1991.

© Andrew Pawley

particular subject matter code – that used by *Wisden cricketers' almanack* for written summaries of cricket matches – appears at the head of this paper.

I suggest that all normal discourse is framed in terms of one subject matter code or another. Certain codes are common to all mature speakers of a language. Others are used mainly or exclusively by a subgroup. When people speak of 'medical language', 'legal jargon', 'criminal argot' and so on, they presumably have in mind such specialised codes. Usually it is easy for someone who knows a subject matter code to tell whether an interlocutor commands the code or not – the latter needs to talk about the subject in question for only a few seconds to give the game away.

That is all very well. But the sixty-four dollar question, for a linguist, is whether subject matter codes are reasonably well defined systems. Are they amenable to rigorous analysis? What conceptual tools do we have for this task? With these concerns in mind, I will examine in some detail a particular subject matter code – the gentle science of radio commentary on games of cricket.<sup>2</sup>

Radio cricket commentaries have much in common with various other traditions of expert speech that are sometimes termed *oral formulaic genres*. The model of analysis I will use owes a good deal to the work of Koenraad Kuiper, of the University of Canterbury, and his associates on radio sporting commentaries and other English formulaic speech genres (see Kuiper and Haggio 1984, 1985, and other references). Each of the speech traditions examined by Kuiper and his co-authors shows five features that (taken together) distinguish them from other discourse genres.<sup>3</sup> The five features are:

- (a) very strict discourse structure rules, specifying the topics proper to the discourse and their order of occurrence;
- (b) a very high concentration of speech formulas, each indexed to a particular discourse context or range of discourse contexts, giving appropriate information about the topics;
- (c) special grammatical rules;
- (d) special prosodic or musical patterns;
- (e) exceptional fluency, i.e. fewer than average unplanned pauses.

Kuiper et al. outline a framework for constructing generative descriptions of such speech traditions. The descriptions are intended to be generative in two senses. First, they seek to be explicit, defining in a precise manner the object of inquiry and its structure. Second, they seek to be predictive, formulating rules for the production of acceptable utterances or texts which go beyond the corpus of recorded examples. No name has been given to this descriptive framework. We might call it *generative formulaic*.

The question arises why formulaic speech genres have characteristics (a)-(e) – what are the discourse conditions or other historical factors that have favoured the development of these features? In this connection what makes radio commentary genres of particular interest is that they have arisen

<sup>2</sup>In deference to George Grace's American origins and his fondness for citing baseball examples in his writings, I should perhaps have taken as my subject baseball commentaries. These have much in common with cricket commentaries. Alas, I am not a native speaker of the language of baseball. But readers should be aware that G.W. Grace, the linguist, shares his surname and both initials with W.G. Grace, the greatest name in nineteenth century cricket.

<sup>3</sup>Kuiper and his co-authors mention only four distinctive elements. However, I notice that a fifth (special grammar) occurs in all the genres of commentary that he and his co-authors have written about.

within living memory and their evolution is fairly well documented. Hence the making of a discourse genre can be studied almost at first hand.

Three main parts or subgenres of radio cricket commentaries can be distinguished, differing from one another in discourse structure, vocabulary, grammar and musical patterns. Two of these (play-by-play descriptions and score summaries) are highly formulaic, the third (colour commentary) less so. My explanations of the characteristics of each subgenre will focus on discourse conditions and I will not attempt to study their historical development.

Finally, I will ask whether the idea of a subject matter code and the generative formulaic descriptive framework can profitably be applied to discourse genres that are not usually considered to be formulaic. I will suggest that, in large part, they can be.

### 1.1 ON GRACE ON WAYS OF TALKING ABOUT THINGS

My notion of a subject matter code is near kin to Grace's *subject-matter-consecrated ways of talking* (Grace 1987:103).

To use a language correctly, Grace (1981, chs. 3-6; 1987, chs. 3, 7) says, one needs to know the speech community's ways of talking about things – what things are sayable and how they are said. But what exactly are these 'things' that Grace regards as the stuff of discourse? He distinguishes two sorts. One – the things people talk about – he equates with the very general notion of *subject matters*. Members of a speech community will develop a body of subjects, topics or themes of discourse that reflect the conceptual worlds and concerns of its members. Speech communities which have markedly divergent cultures will have rather different sets of conventional subject matters.

But people also say particular things. When Grace talks of someone saying something he gives to 'something' a quite specific sense. To say something, in Grace's sense, a speaker must *specify a conceptual event or situation*, give the specification a *modality* and *contextualise* it. That is, the speaker must form a construction in which particular constituents correspond to particular parts of an event or situation (the actor, the action, the place, etc.), he must indicate whether the construction is being asserted, denied, commanded, questioned, speculated about, etc. and he must connect the text to the discourse context.

In most languages the minimal grammatical apparatus needed to say something is that of the simple sentence or clause. This correlation between clause grammar and the semantic structure of events and situations seems to be a basic feature of human language design. (I will henceforth use 'event' for both event and situation.)

However, saying things that makes sense to the members of a speech community is not just a matter of producing utterances that are grammatical and make sense in context. We also need to speak *idiomatically*, that is, to say things the way native speakers do. Grace (1981:41,46-54; 1987:93-94,105) mentions several kinds of situations which remind us that there is more to speaking a language than just knowing the meanings of individual words and the rules of sentence formation. One such situation is when we come across a text produced by a foreigner that is perfectly grammatical but quite unidiomatic. Another is when, armed with a good dictionary and grammar book, we are unable to make sense of a piece of text in an exotic language. But we don't need to go so far abroad to find examples. It is notorious that people who work in the same field, and who think of themselves as speaking the same language, at times have difficulty understanding what each other

is saying. In the sciences, for example, fundamental misreadings of texts (or of the writer's or speaker's intended meaning) are fairly common.

Why should this be so? One reason is that even the most descriptive things we say, the 'just-the-bare-facts' reports of events and situations, are of necessity caricatures, extremely simplified sketches or interpretations. When people utter a sentence that purports to describe an event they have witnessed, they do not (and cannot) describe everything that happened in an objective and exhaustive manner. They must *put a construction* on what they observed. The first construal is cognitive: the observer notices certain details and makes sense of them.

The second construal is linguistic: the speaker gives names and grammatical status to certain details, characterising these as objects or processes, and so on, and specifies particular semantic/grammatical relationships as holding between these named entities. Even the words and phrases that name particular details are themselves loaded choices, in so far as each word or phrase is picked from a range of roughly synonymous options each having somewhat different senses. When the conceptual constructions people are talking about are highly abstract, as is the case in much scientific discourse, the possibilities for confusion are considerable.

Given that linguistic representations of everyday events are so subjective and sketchy, why aren't misunderstandings between speakers of the same language even more common than they are? One reason, I think, is that the members of a speech community learn similar conventions for making cultural sense of the world. They put this shared knowledge and belief (call this *cultural* or *backing knowledge*) to work in identifying and talking about entities and events. A kind of shorthand develops for talking about recurring events, allowing discourse to be streamlined for speed. From constant practice in listening to and formulating talk about particular subjects, neophytes learn what kinds of details are regarded as significant, what it is important to mention, what should or may be omitted, and what kinds of information not explicitly stated in the discourse of others can be inferred. Thus, if I ask you what you did last night in London and you answer '*I went to the movies*', I will infer not only that you travelled to a film theatre but also that you obtained a ticket to see the film, went into the theatre, sat down in a seat and watched a feature film which probably lasted for the better part of two hours – because these are things that are normally part of the event *Going to the movies*. One part of this event sequence has come to stand for the whole.

To the layman the most conspicuous units in this shorthand system are single words. But in fact the larger part of the vocabulary of discourse probably consists of phrases and clause-sized expressions that stand for complex conventional concepts. Call these the 'phrasal lexicon'.<sup>4</sup> If you know the words used to talk about a particular subject matter but not the phrasal lexicon you will have a hard time following discourse about that subject matter. If you also lack the backing knowledge needed to infer the missing standard concomitants of event-descriptions, you will be completely lost.

As an illustration of the kind of backing knowledge needed to make sense of a single clause of cricketing discourse and of the great economies of expression made possible by a technical lexicon, consider this short extract from the *Wisden* match report quoted above:

---

<sup>4</sup>After Becker (1975:60).



- (1) Robinson made certain of a lengthy lead on first innings with a determined maiden Test hundred...Most of Robinson's runs came off his legs and behind square on the off side, where by opening the blade he repeatedly found gaps.

An accurate translation of (1) into non-technical English is, for all, practical purposes, impossible. The best that can be done is to provide a gloss or paraphrase with extensive explanatory notes. The gloss and explanations should draw on at least the following information about how cricket is played (additional information, not essential in order to gloss (1) but likely to be useful in interpreting later texts, is added in parentheses):

- (a) Games of cricket are played on a large (oval or circular) *field* (ideally, about 150 yards across) between two *teams* or *sides* (of 11 players each).
- (b) Games between international teams are called *Test matches* or *Tests*.
- (c) At any time in a game one team *bats* while the other team *fields*. The batting team (and each of its members) is said to *have an innings*.
- (d) Play is started when one player in the fielding team, called the *bowler*, delivers a ball to another player, called the *batsman*. Each such act is called a *ball* or *delivery*. The batsman has a bat for striking the ball in order to *defend his wicket* or to *score runs*.
- (e) The wicket consists of three cylindrical sticks, called *stumps*, with two small cross-pieces, called *bails*, resting on top. The wicket (28 inches (71cm) high and nine inches (23cm) across) is said to be *broken* or *down* if a bail is dislodged.
- (f) A bowler delivers the ball by *bowling* it (i.e. without bending the elbow at moment of release), usually overarm, from beside another wicket (set 22 yards away from the batsman's wicket). (The surface between the wickets, is usually turf; this surface, called the *pitch*, should be close-mown, and flattened and hardened by repeated heavy rolling.)
- (g) The fielding team's aim is to get the batsman *out* and to prevent him from scoring runs. (A batsman may be got out in various ways, e.g. if the bowler knocks his wicket down with a delivery; if the batsman hits the ball and it is caught before it hits the ground by a fielder; or if when a fielder knocks the wicket down with the ball the batsman is out of his ground or safe area, which is marked by a line across the pitch about four feet (122cm) in front of the wicket called the *popping crease*.)
- (h) A bowler usually aims to bowl in line with the stumps and to pitch the ball in front of the batsman so that it is rising when it reaches him. (Bowlers alternate. One bowler bowls a series of six deliveries in succession, called an *over*, from one *end* of the pitch, then another bowler bowls an over from the other end, and so on until the end of the innings.)
- (i) Two members of the batting side are always on the field. During any one delivery one of these batsmen receives the bowling and is said to be *on strike* or *the striker* while the other (the non-striker) waits by the bowler's wicket.
- (j) A batsman scores one run if after hitting the ball with the bat he and his batting partner each runs to the opposite wicket before the fielding side can gather the ball and knock down either wicket with it. (Further runs may be scored if the pair continue to run back and forth between the wickets before the fielding side are able to break the wicket. If the



- hit rolls or bounces over the boundary of the playing field four runs are credited to the batsman. If the hit clears the boundary on the full six runs are credited.)
- (k) Runs can be scored by hitting the ball to any part of the field. (A batsman does not have to attempt a run if he hits the ball. The batsmen usually choose to run only when the striker hits the ball a safe distance from any fielder.)
  - (l) A batsman continues his innings, receiving deliveries, until the fielding team gets him out. There is in principle no limit to the number of runs a batsman may score in one innings but it is a notable feat to score 100 runs. A score of 100 runs or more by one batsman in a single innings is called a *century* or a *hundred*.
  - (m) When a batsman gets out he is replaced by another member of his team, until all but one of the (11) members of the batting side are out, at which point the batting team has completed its innings. (The total score for the team's innings is computed by adding all the individual batsman's run scores plus extra runs due to illegal deliveries, etc.) At that point the other team takes an innings.
  - (n) In Test matches each team is entitled to two innings each, taken alternately. The side with the highest score after its first innings is said to *lead on the first innings* or to *have a first innings lead*.
  - (o) The three stumps in the batsman's wicket are called the *leg stump*, the *middle stump* and the *off stump*.
  - (p) At the moment of delivery the batsman stands in front of his wicket, usually straddling the popping crease. He usually stands side-on to the bowler, with his feet placed on the leg stump side of his wicket, giving the bowler a clear view of at least the middle and off stumps.
  - (q) The field has two sides, defined in relation to an imaginary line running through the two sets of wickets and in relation to the batsman's stance. The *leg side* or *on side* is that half of the field on the leg stump side of the batsman. The *off side* is the other side.
  - (r) The blade of the bat has a flat front surface, called the *face* (4.25 inches or 10.80cm wide), which helps the batsman to control precisely the direction of the stroke.<sup>5</sup> By driving straight through the line of the ball with *full face*, the batsman may hit the ball straight in front of him. By *closing the face* of the bat at moment of impact, the batsman may direct the ball to the leg side. By *opening the face* of the bat at moment of impact he may direct it to the off side.
  - (s) The bowler is supported by (10 other) fielders placed at various positions around the field. (One of these fielders, called the *wicket-keeper* or *keeper*, stands behind the batsman's wicket to catch deliveries which pass the bat or come off the edge of the bat.)

<sup>5</sup>It is primarily the flat face of the bat that makes batting in cricket a more exact science than batting in baseball. The possibility of hitting to all parts of the field, behind as well as in front of the batsman, also allows for a greater variety of strokes. I suggest also that there are two technical differences that make bowling in cricket a more varied art than pitching in baseball (though the latter is a considerable art). One is the fact that in cricket the ball normally hits the ground before reaching the batsman. Deliveries which reach the batsman on the full are part of the bowler's repertoire, but the ball is normally bowled so as to pitch a few yards in front of the batsman and the bowler seeks to create difficulties by means of variable length, bounce and break as well as by the initial flight or swerve and direction of the delivery. The other main difference is that a cricket ball has a raised seam of stitches and (when new) a shiny surface. Bowlers are able to exploit these attributes to obtain swerve in flight and break or deviation off the ground after the ball pitches.

- (t) One speaks of any position or trajectory as being *square of the wicket* if it is on an imaginary straight line running across the field intersecting the three stumps of the batsman's wicket. Any point in front of this line, i.e. towards the bowler's end, is *in front of the wicket*; any point behind this line is *behind the wicket*. The further a point is from square of the wicket the *finer* it is. This terminology is used to distinguish both fielding positions and the direction of strokes.

The following gloss, to be read in conjunction with notes (a)-(t), is now offered for (1):

- (2) The batsman Robinson scored more than 100 runs in a single innings in a Test match for the first time and this high score ensured that his team's total of runs in its first innings was much larger than the other team's first innings total. Robinson scored most of his runs in two ways. He scored many from deliveries which reached him on or about the line of the leg stump and which were pitched close to him, so that he was able to hit the ball on the half-volley by closing the face of the bat, thus angling the ball square or behind square on the legside. He also scored many runs from deliveries that reached him just outside the line of the off stump, allowing him room to move across and open the face of the bat as it struck the delivery, so that the ball was deflected behind square on the off side, speeding through gaps between the fielder placed directly square of the wicket on that side and those fielders placed in finer positions behind square.

Such is the extent of the background knowledge that we take for granted in specialist discourse that even with these explanatory notes (1) will no doubt be only partially intelligible to many readers. To gloss the complete extract from the Wisden match report, cited at the head of the article, would take many more explanatory notes.

Grace (1987:32-33,100) makes the point that you cannot completely separate *what* is said (the meaning) from *the way* it is said (the form). However, he also acknowledges that in treating ordinary language it is necessary to distinguish a third level of saying things which depends on just such a separation. We recognise certain sets of diverse expressions or utterances as being in some sense equivalent even though their literal meanings are different. For instance, the phrases *I'm very sorry!*, *My apologies!* and *Please excuse me!* can all serve, in certain contexts, as an apology. Similarly, *'The Giants beat the Dodgers easily in game one'* and *'The Dodgers were thrashed by the Giants in the opener'* give the same essential information about the outcome of a particular game of baseball.<sup>6</sup> While these different expressions do not *literally* say the same thing, they do say *essentially* the same thing at a certain level of abstraction – one where identity and contrast of concepts must be defined in terms of discourse functions in a particular subject matter code. I will speak of such utterances as being *functional equivalents* in a given discourse frame or context. Grace's (1981:92,102) "paraphrase set" represents the same idea.

<sup>6</sup>Many American sports broadcasters use the formula '*X defeated Y*' when announcing the results of a series of games, but follow the convention of using a different expression for '*defeat*' in announcing each result. It is noteworthy that among the long list of terms they draw on (e.g. *demolish, blast, topple, edge, outlast, squeak by*) are many which do not literally mean '*defeat*'. Their interpretation as having this pragmatic meaning depends on the listener knowing the discourse frame.

## 2. RADIO CRICKET COMMENTARIES

Now let us turn to radio cricket commentaries.<sup>7</sup> A competent cricket commentator must possess (at least) four types of expertise: (i) an extensive knowledge of how cricket is played, that is, of the rules and procedures of the game, (ii) trained perceptions that enable one to pick out the significant details of events and situations while watching a particular game, (iii) knowledge in principle of how to formulate an account of these details in a way that cricket followers will readily understand, and (iv) the ability to create high quality commentary under time pressure, drawing on (i)-(iii) to formulate descriptions and analyses that are not only clear but entertaining.

Commentators take it for granted that their listeners are steeped in cricket culture and know a great deal about (i)-(iii). Without such knowledge and skills the audience could not make sense of the commentaries. What distinguishes good commentators, chiefly, is their skill in (iv), the creation of high quality commentaries.

### 2.1 BALL-BY-BALL COMMENTARIES

In the eight major cricket-playing countries<sup>8</sup> three main classes of cricket matches are regularly broadcast on radio:<sup>9</sup> (a) international matches played over five or six days (these are called Test matches), (b) other matches played over three or four days between regional or other teams of recognised 'first-class' status, and (c) one day matches or 'limited overs' played by international and certain other teams. Only matches of types (a) and (b) have official first-class status. The rules of play for type (c) matches differ slightly from those for first-class matches.

Radio commentaries are often broadcast for the full duration of a match. A day's play lasts for around six hours and in first-class matches is divided into three main 'sessions' or 'periods' of about two hours each: a morning or pre-lunch session, an afternoon or post-lunch session and an evening or post-tea session. While play is in progress it is customary for two commentators to work together for spells of 20 minutes or so, combining to do what in the trade is called *ball-by-ball commentary*. The term refers to the coverage of each minimal unit of play, termed a *ball* or *delivery* because play is officially started when the bowler runs in to bowl or deliver the ball to the batsman on strike. The end of that piece of play occurs when the ball has become *dead*. The ball becomes dead when a batsman has been given out, or when the ball has crossed the boundary line, or when it has been gathered by the fielding side and the batsmen have clearly given up trying to take runs.

<sup>7</sup>Television commentaries on cricket are quite different from those done on radio. Camera-work often substitutes for descriptions of play, scores and scenes. In fact, experienced TV commentators do very little straight descriptive play-by-play commentary. They assume that the viewers know the game and will prefer to 'read' the pictures themselves.

<sup>8</sup>The eight nations that have played Test cricket are England, Australia, South Africa, West Indies, New Zealand, India, Pakistan and Sri Lanka. (The West Indies team is actually drawn from several Caribbean nations.) My impression is that commentators from all these countries use the same code, with very minor regional variations.

<sup>9</sup>According to the English broadcasters John Snagge (1962) and Peter West (1962) the first ball-by-ball commentaries were broadcast by the BBC in about 1935. In the 1920s and early 1930s eyewitness summaries of periods of play were given, but in those early days of radio it was thought that cricket was too slow for a live commentary on it to hold the attention of an audience.

## 2.2 DISCOURSE GENRES WITHIN BALL-BY-BALL COMMENTARIES

A ball-by-ball commentary itself consists of at least three distinct discourse genres which I will refer to as *play-by-play descriptions*, *score summaries* and *colour commentary*.<sup>10</sup>

Play-by-play descriptions (calling the action as it happens) and score summaries both exhibit the several core characteristics of formulaic speech, to be discussed in more detail below.

The term colour commentary subsumes all remarks made by either of the commentators in the intervals between play-by-play accounts and score summaries. Broadcasters use these intervals to speak their minds on various matters, including but not limited to the events of the last delivery. In the relative flexibility and diversity of its subject matter and discourse structure, colour commentary is fairly typical of ordinary, conversational talk.

During any one spell of broadcasting, the two commentators do complementary jobs. One, whom I will call the *principal commentator*, handles the more demanding work of play-by-play description. The principal commentator also usually provides a first recapitulation of the key events in each piece of play after the play-by-play description is done and does the score summaries. The other, whom I will call the *colour commentator*, shares with the principal commentator the job of talking about the game between deliveries.

Transcripts of some typical segments of commentary follow.<sup>11</sup> The letters P, C and S in the margin represent the start of a play-by-play segment, a colour commentary segment, and a score summary, respectively. The numbers beginning each line of transcript indicate subject matter units, whose functions will be identified below, in Table 1. Full stop and comma mark contour final and non-final intonation. Pauses are marked by hyphens, - representing 0.2 to 0.5 of a second, -- representing 0.6 to 1.0 and --- a pause longer than one second. CAPITALS mark loud speech. Italics mark speech that has most or all of the bundle of prosodic features defined below as 'dramatic mode'.

- (3) P 1 *McPhee, -- in, -*  
 4 *to Waugh, --*  
 6 *THIS time, -- short! -*  
 8 *Waugh again just glides it*  
 9 *down towards third man. -*
- C That shot's been repeated two or three times this over. -  
 A single taken on each occasion from both Tucker and Waugh. -
- S Waugh one hundred and ninety TWO. -  
 Six for four hundred and twenty one. ---
- C Time for one more over after this one. -  
 Waugh showing no signs of panic, -  
 just continuing to play the way he's played all day.

<sup>10</sup>The first and last of these terms are not, as far as I know, used by cricket commentators. I have borrowed them from Kuiper and Haggio's (1985) analysis of ice-hockey commentaries.

<sup>11</sup>The data used for this analysis were taken from broadcasts of the following matches: Australia v. Pakistan (ABC, 25 February 1990), Australia v. New Zealand (ABC and RNZ, 15-19 March 1990) and NSW v. Tasmania (ABC 10 March 1990).

- (4) P 1 *As we watch, - COOley- steaming in now, -*  
 3 *bowls*  
 4 *to Waugh again, -*  
 8 *just stroking it*  
 9 *out on the on side, -*  
 16 *thinking about a single, -*  
 16 *Tucker taking a few ah - stuttering steps down the wicket from the bowler's end*  
 17 *but Waugh sending him back. -*  
 C There was no run there, as the ball was played straight to the fieldsman. -  
 S Six for four hundred and seven.  
 Mark Waugh one hundred and eighty five, -  
 Tucker on fourteen.

In the next example it is hard to decide where the play-by-play report ends and the colour commentary begins, because the commentator begins to recapitulate (at 'that ball was VErY short inDEED') before he has finished describing the action.

- (5) P 4 *Tucker on strike now, -*  
 C He's on fifteen, -  
 No doubt he'll be keen to feed the strike to Mark Waugh. -  
 P 1 *But Cooley,*  
 2 *coming to the end of this over*  
 3 *as he bowls,*  
 4 *and Tucker, -*  
 6 *finds a SHORT ball,*  
 5 *NO BALL in fact, called,*  
 6 *down the leg side. -*  
 6 *That ball was VErY short inDEED. -*  
 6 *and also well wide of the leg stump.-*  
 7 *Tucker - moved inside of it*  
 8 *but really didn't play a shot at the ball,*  
 12 *as Soule taking the ball going down the leg side. ---*  
 S Sundries now thirty one, --  
 Six for four hundred and twelve. -  
 The lead quickly approaching three hundred.

I now turn to a detailed examination of the first of these three domains of commentary.

### 3. PLAY-BY-PLAY DESCRIPTIONS

The question of how commentators go about formulating play-by-play descriptions may be broken down into four: Why, what, when and how? That is to say: What is the purpose of the commentators? What is their agenda (what bits of information do they supply to the listener and in what order)? And how is the information presented? The reader who would point out that these four questions can usefully be asked about the construction of discourse in *any* subject matter has anticipated our conclusions.



## 3.1 PURPOSE, SUBJECT MATTER AND CONTEXT

Play-by-play descriptions are typically brief but dense. The broadcaster's main purpose in these is to communicate to the listener the drama of the moment by calling the action as it happens, or as near to it as possible. A complete play-by-play segment usually takes between five and fifteen seconds, depending on what happens during a particular ball or delivery (i.e. the period where the ball is 'in play') and on what the commentator makes of these happenings. As a rule there follows immediately a more leisurely review or recapitulation of events during the delivery. This review is better treated as part of colour commentary rather than as part of the play-by-play description.

During each ball a number of things will happen that an audience knowledgeable about the game will judge to be significant and want to be informed about. It is these *significant* or *comment-worthy events* that make up the potential subject matter of a play-by-play description. During moments of action, then, the principal commentator must watch the play closely, spot the significant details and describe as many of them as can comfortably be fitted into the commentary without falling far behind the action.

Table 1 is an attempt to distinguish the main types of event-types which were mentioned in a sample of 100 play-by-play descriptions and it lists these in the order in which they were most often given.<sup>12</sup>

TABLE 1: THE SUBJECT MATTERS OF SIGNIFICANT EVENTS REPORTED  
IN PLAY-BY-PLAY DESCRIPTIONS

(Event-types 1-6 are concerned with the act of delivery)

1. BOWLER'S NAME AND APPROACH
2. BALL OF THE OVER (first ball, second ball, etc.)
3. MOMENT OF DELIVERY
4. NAME OF BATSMAN ON STRIKE
5. FAIRNESS OF DELIVERY (call or no call of 'no-ball' or 'wide')
6. DESCRIPTION OF DELIVERY (pace, length, line, lateral movement, etc.)

(Event-types 7-12 are concerned with the batsman's response)

7. BATSMAN'S FIRST MOVEMENT (back, forward, etc.)
8. BATSMAN'S STROKE/NO STROKE (description of the attempted stroke or leave)
9. DIRECTION OF STROKE
10. HEIGHT OF STROKE
11. DISTANCE OF STROKE

<sup>12</sup>A case can be made for including the bowler's walk back to his mark as the first potential topic in play-by-play descriptions, because commentators sometimes mention this activity and occasionally even remark on the manner of the walk-back. I prefer to treat such remarks as part of colour commentary. Here I have treated the bowler's run up to the wicket as the first element in play-by-play reports because in my material this is consistently the first event described in dramatic mode. There is a happy coincidence with the rules of cricket, which specify that the ball is in play once the bowler *begins* his run to the wicket but not before.

I have also excluded other topics that are occasionally interwoven with the bowler's approach. These include the bowler's field-placings ('*Bracewell, bowling to a six-three field, comes in to Border*'), the time ('*Bedser moves in, with the clock now showing one minute to six*'), the weather ('*Trueman begins his run, with light rain falling at the Oval*') and crowd behaviour ('*The crowd chants "Had-lee!" as he sets off again*'). Most of these items have the character of colour comments, sneaked in at the beginning of the play-by-play report. However, it might be argued that mention of the bowler's field-placings belong as much to the play-by-play report as identifying which ball of the over is being bowled.



(Event-types 12-18 are concerned with the outcome of infield action)

12. IMMEDIATE OUTCOME IF STROKE IS UNSUCCESSFUL

Any one of:

- a. NO CONTACT (batsman plays at ball but misses)
- b. BAT CONTACT (the ball strikes the bat)
- c. BODY CONTACT (the ball hits the batsman's pads or body)
- d. BOWLED OUT (the ball hits the stumps and dislodges one or both bails)

13. FIELDING CLOSE TO THE WICKET:

Any one of:

- a. CATCH (named fielder/fielding position catches the batsman out)
  - b. MISS (named fielder/fielding position misses a catch)
  - c. STOP (named fielder/fielding position stops the ball)
  - d. STUMPING ATTEMPT (wicket-keeper gathers ball and attempts to break wicket with the batsman out of his ground)
  - e. BALL RUNS AWAY FOR POSSIBLE BYES OR LEG-BYES
14. APPEAL (for catch, stumped, leg-before-wicket, etc.)
15. UMPIRE'S DECISION (batsman is out or not out)
16. BATSMAN'S CALL (call/no call)
17. PARTNER'S RESPONSE (acceptance/refusal/confusion)
18. RUN TAKEN / NOT TAKEN (for hit, byes, leg-byes)

(Event-types 19-27 are concerned with the outcome of outfield action)

19. CHASE OR WAIT FOR CATCH BY NAMED FIELDER
20. GATHER /MISS (fielder makes a catch/stop/pick-up or misses the ball)
21. APPEAL FOR A CATCH
22. UMPIRE'S DECISION
23. THROW-IN (description of fielder's return throw)
24. FURTHER RUNNING (batsmen take/decline further runs)
25. OUTCOME OF THROW-IN (ball is taken by keeper/hits stumps/goes for overthrows)
26. APPEAL FOR RUN OUT
27. UMPIRE'S DECISION ON RUN-OUT APPEAL

[IF RUNS FOR OVERTHROWS ARE ATTEMPTED START COMMENTARY AGAIN AT EVENT 16 ]

It is worth emphasising that there a finite number of topics that it is proper to mention in a play-by-play report and that Table 1 comes close to listing them all. However, although Table 1 is the first step towards a formal description of play-by-play reports it leaves out much essential information. It does not say how particular event-descriptions may be worded. Nor does it show all the possible combinations and permutations of event-types.

Let us now look at a few commentary extracts and see how they square up with Table 1. Examples (6)-(11) illustrate typical minimal play-by-play descriptions, consisting of four to six significant events reported in the space of three or four clauses. In these and later examples any fragment of the commentary that describes a single event is given a separate line, regardless of

whether that fragment is a full clause or not. The numbers preceding each line in the transcripts indicate that the subject matter(s) correspond to the categories numbered in Table 1.

- (6) 3, 2 Tauseef starts a new over, -  
 8 *Moody deflects,* -  
 9,11 *down to third man.* -  
 18 Can't get much easier runs than that.
- (7) 3 *Tauseef bowls*  
 4 *to Border,* -  
 12 *strikes him on the pad,* -  
 14 *Tauseef THROWS his arms up in appeal!* --  
 15 The umpire isn't interested.
- (8) 1 *Aaqib runs in,* -  
 8 *O'Donnell pushes*  
 9 *towards point,* -  
 16 *they look for a single,* -  
 18 *they don't go,* -  
 20 Ijaz picks up.
- (9) 3 *Aaqib*  
 4 *to Border,* -  
 6 *and Border gets a ball that again kept low*  
 8 *and he pushes it*  
 9,11 *out to mid-off.*
- (10) 1 *Away goes Aaqib,* -  
 3 *and delivers*  
 4 *to Border again,* -  
 8 *Border turns him*  
 9,11 *down to fine leg.* -  
 18 The batsmen take a single.
- (11) 1 *Waqar Younis, full of life, runs in,* -  
 8 *Moody plays it*  
 9,11 *back to him,* -  
 18 and the batsmen stay at home.

It can be seen that the most commonly mentioned events in these short descriptions are 1 (bowler's name and approach) or 3 (delivery) (not both), 4 (naming the batsman on strike), 8 (batsman's stroke), 9 and 11 (direction and distance of stroke) and 18 (whether or not a run was taken).

It is however common to find as many as seven or eight events receiving comment in a single play-by-play description of 10 seconds or so. Examples (12)-(20) illustrate descriptions which refer to a range of other events besides or instead of those which recur in (6)-(11).

- (12) 3 *Waqar bowls*  
 4 *to Taylor, -*  
 7 *Taylor is back, -*  
 8 *he turns*  
 9 *around the corner, -*  
 19 *Akram gives chase, --*  
 18 *the batsmen take a single.*
- (13) 3 *Waqar bowls*  
 4 *to Jones. -*  
 8 *Jones drives him*  
 9 *straight! -*  
 11 *That'll go for FOUR! -*  
 19 *X is chasing, -*  
 24 *They take in TWO, -*  
 24 *they turn for THREE, -*  
 23 *as the return comes in.*
- (14) 1 *In comes Ghouri, -*  
 8 *Taylor hits this*  
 10 *in the air,*  
 11 *safe enough, -*  
 9 *over mid-wicket. -*  
 20 *X is there, -*  
 18 *Taylor takes a single.*
- (15) 3 *Naadem bowls again*  
 4 *to Moody, -*  
 7 *he's down the pitch, -*  
 12 *MISSES! -*  
 13 *Is he stumped THIS time?*  
 15 *He's OUT I think! -*  
 13 *He's STUMPED by Younis.*
- (16) 3 *Mustaq bowls*  
 4 *to Jones,-*  
 12 *and Jones is BEATEN! -*  
 14 *There's a LOUD APPEAL from Mustaq! -*  
 15 *NOT OUT! -*  
 15 *Very close! -*  
 14 *Mustaq has a second bite at the umpire.*
- (17) 1 *Here comes - MacPhee, -*  
 8 *DRIVEN BY WAUGH!*  
 8 *THAT'S a good stroke! -*  
 9 *THROUGH the covers!*  
 11 *It won't go all the way to the boundary*  
 18 *They're coming back for TWO. -*

- 19 *and it's Cox, - Jamie Cox who chases back. -*  
 22 *Good throw to Soule.*
- (18) 3 *Akram bowls again, -*  
 8 *and Taylor cuts,*  
 7 *he leans back*  
 8 *and hits that one*  
 9 *through point. -*  
 8 *A lovely shot! -*  
 11 *It won't reach the boundary unfortunately,*  
 20 *Rameez is there. -*  
 18 *They only take a single.*
- (19) 3 *Mustaq bowls*  
 4 *to Jones, -*  
 7 *down the pitch, Jones, -*  
 10 *slightly uppishly,*  
 9 *down the ground, -*  
 20 *Aaqib fields, -*  
 18 *the batsmen take a single.*
- (20) 3 *Mustaq bowls, -*  
 8 *Moody drives hard*  
 9,11 *to mid-off. -*  
 18 *Only for a single.-*  
 20 *Fielded by Shoaib Mohammed, -*  
 21 *who breaks the stumps with his return.*

Table 2 gives the frequency of mention, in a sample of 50 play-by-play descriptions, of nine event-types which are constituents of every normal delivery.

TABLE 2: FREQUENCY OF MENTION OF PARTICULAR EVENT-TYPES IN 50 PLAY-BY-PLAY DESCRIPTIONS

	Times Mentioned	Percentage
1. Bowler's approach	21	42
2. Ball of the over	14	28
3. Moment of delivery	39	78
2 + 3: Approach and/or delivery	50	100
4. Batsman on strike	33	66
5. Fairness/unfairness of delivery	2	4
6. Details of delivery	10	20
7. Batsman's first movement	15	30
8. Batsman's stroke/no stroke	47	94
9. Immediate outcome of delivery if stroke unsuccessful	12	24
8 + 9: Outcome of delivery	50	100

Space limitations prevent a detailed discussion of how commentators handle all the potential events in a delivery. Brief comments on a number of events follow.

## BALL OF THE OVER

It is not obligatory to specify whether the current delivery is the first, second, etc. ball of the sequence of (normally) six that make up an over. The likelihood of a commentator specifying which ball of the over it is depends on its place in the sequence. If a ball is the first or the last of the over it is usual to mention the fact, as for example, in commentaries (5) and (6). Other positions in the over seldom get a mention.

## BOWLER'S NAME AND APPROACH

Whether the bowler's approach will be reported depends largely on the length of his run to the wicket. *Slow* or *spin* bowlers (those who rely primarily on flight and spin) usually amble up to the wicket from 5-8 yards and their run-ups were mentioned in just 17 per cent of cases. *Pace* or *fast* bowlers (those who bowl at fast and medium-fast pace) take a long and often spectacular run-up, normally between 10 and 15 accelerating paces. The approach of pace bowlers was reported in 59 per cent of the commentaries. Commentaries (4) and (11) contain little sketches of a pace bowler's run-up, as does the following:

- (21) S Tucker on sixteen.  
 P 4 *He's on strike. -*  
 1 *McPhee, - slowly moves in now, -*  
 1 *gathers pace, -*  
 3 *and bowls to Tucker, -*  
 6 *outside the off stump*  
 7 *and plays this*  
 9,11 *down to third man...*  
 18 One more to Tucker,...

As part of the description of the bowler's approach the commentator may mention the end from which he is running, for example, '*Botham comes in from the Vulcan Street end*'.

## ACT OF DELIVERY

There are two kinds of information that are provided in every play-by-play description: (i) the act of delivery, (ii) the immediate outcome of the delivery. These correspond to the two 100% lines in Table 2. It is understood that (i) should be announced as it occurs, and (ii) as soon as it has occurred. A commentator who is late with these details will apologise.

The act of delivery is signalled by a combination of words and prosodic features. The delivery itself may be announced (a) indirectly, by mentioning the bowler's approach to the wicket, followed by an account of the batsman's response, (b) directly, for example, by the formula *X bowls to Y*, or (c) by mentioning both the approach and the delivery, as in examples (8) and (10). Much less often commentators indicate the act of delivery by other means, as in '*X [the batsman] waits on Y [the bowler]*' or '*The next ball is driven by Wright...*' The precise moment of delivery is usually signalled by the pause pattern and tempo of speech. Up to the point of delivery the commentator's rate of articulation is leisurely and the volume moderate. There is sometimes a significant pause in the moment before the operative words announcing the release of the ball, as, for example, in (3), (5) and (22). Often this slow build-up is followed by a sudden increase in tempo and volume as the batsman plays the ball.

#### NAME OF BATSMAN ON STRIKE

The batsman's name is often worked into the clause that describes the bowler's delivery, as, for example, in (3), (4), (7), (10), (12), (13), (15), (16) and (19). Just as often, however, it is delayed until the description of the batsman's stroke or first movement, as in (5), (6), (8), (11), (14), (17), (18) and (20). If the name is mentioned along with the delivery it is usually left out or replaced by a pronoun in the account of the stroke but is sometimes repeated, as in (12), (13), (16) and (19).

#### FAIRNESS OF DELIVERY

If a delivery is called a no-ball (i.e. illegal in manner) or wide (out of the batsman's reach) by one of the umpires, the commentator is bound to report the call, as in (5). If a delivery is fair, however, that fact is rarely mentioned.

#### BEHAVIOUR OF DELIVERY

The time between the release of the ball and its arrival to the batsman is less than a second – in the case of really fast bowling about 0.4 of a second. That partly explains why in play-by-play accounts commentators only occasionally (in about 20% of cases) give information about event 6, the *behaviour* of the delivery – for example, about whether it is swinging away from or into the batsman, pitched on a good length, short-pitched, over-pitched, directed at or outside the line of the stumps, breaking sharply after pitching and so forth – separately from their description of the batsman's stroke. (Examples (5) and (9) are exceptions.) There is hardly time for the commentator to mention such details before the stroke is played. Some information about the line and length of the ball is often worked into the account of the batsman's stroke, for example, '*Gower flashes at that outside off stump*', '*Taylor whips it off his pads*' (indicating a ball on the line of the leg stump, probably fairly full in length), '*X flicks it off his hip*' (indicating a fairly high-bouncing ball on the leg stump), '*Walters picks it up on the drop-kick*' (indicating an overpitched ball on the leg stump), or '*Bracewell [the bowler] brings Border [the batsman] forward*' (indicating a fairly straight ball pitching on good-length). However, the place where details of the delivery are most likely to be elaborated is in the recapitulation of events which usually follows a play-by-play commentary.

#### BATSMAN'S FIRST MOVEMENT

Before the ball reaches him the batsman usually makes a rapid adjustment of his feet, often moving back half a pace or moving his front foot forward or even taking one or two full paces towards the bowler. In about a quarter of the play-by-play reports, as, for example, in (12), (15), (18) and (19), such movements are reported independently of the description of the stroke, that is, as event 7. More often information about footwork can be inferred from the description of the batsman's stroke, because each of the major categories of stroke (see list in next paragraph) has its own characteristic footwork.

#### STROKE/ NO STROKE

Table 2 records that the commentators explicitly mentioned in 47 out of 50 descriptions that the batsman played/did not play a stroke. (A deliberate no stroke normally consists either of letting the delivery pass harmlessly through to the wicket-keeper or using the pad on the leading leg to deflect a ball that is pitched outside the off stump.) For all three exceptions there is a simple explanation. In one case the ball was called wide, (i.e. out of the batsman's reach). The other two cases were balls that got the batsman out. In one the audience was told that the batsman was bowled out, in the other that he was stumped, and the listener could deduce both times that the batsman had played at the ball and missed it. A characterisation of the stroke is almost always provided, directly or indirectly. A



direct characterisation is one which explicitly describes the production of the stroke. Often this is done by using a standard descriptive term for a type of stroke, such as *drive*, *cover-drive*, *on-drive*, *chip*, *cut*, *square-cut*, *hook*, *pull*, *sweep*, *reverse sweep*, *glance*, *glide*, *whip off the pads*, *flick off the hip*, and so forth, as in (10), (13), (17), (18) and (20). An indirect characterisation uses a transitive verb like *play*, *stroke*, *hit*, *work* or *push (the ball)*, which tells little about stroke production, in combination with a description of the direction and force of the stroke, which allows the listener to infer the type of stroke played, as in (4), (8), (9), (11), (14) and (21).

#### DIRECTION, FORCE, HEIGHT AND DISTANCE OF THE STROKE

The direction of the stroke is usually given in the same clause that initially reports the stroke, as is the case in (3), (4), (6), (8), (9), (12) and (14). Information about the force, height and/or distance is sometimes also incorporated in the same clause. However, any one of these features may rate a separate clause. In particular the distance travelled is usually described separately, as in (13), (17), (18) and (22).

#### RUN TAKEN / NOT TAKEN

If a run is taken the fact must be mentioned. The absence of a run may be signalled either by an overt statement, such as '*No run there*' or '*The batsmen stay at home*', or simply by no mention of running. In the latter case, other information will have been given that allows the listener to infer that no run was taken, for example, the commentator will say that the batsman has hit straight to a close-in fielder, or has let the ball go through to the wicket-keeper.

#### IMMEDIATE OUTCOME IF STROKE UNSUCCESSFUL

If the stroke or leave is unsuccessful, the fact must be reported. The most common unsuccessful outcomes are: (a) the batsman plays and misses a ball that goes through to the wicket-keeper, (b) the ball strikes the edge of the bat or is otherwise mishit, (c) the batsman is hit on the body, usually on the pads that cover each leg, (d) the batsman's stumps are hit by the delivery -- he is bowled out.

### 3.2 RESIDUAL PROBLEMS IN DISCOURSE STRUCTURE ANALYSIS

A full account of the discourse structure rules of play-by-play descriptions would take more space than we are allowed. Description is complicated by the fact that some events do not occur in every delivery but must be mentioned when they do, while some other events occur in every delivery but need not be mentioned.

It would be instructive to compare how similarly different commentators report the same deliveries. On this point I have only casual observations, obtained by comparing simultaneous radio and TV commentaries, and by listening to the principal commentator and the colour commentator compare notes. My impression is that, in spite of the subjective nature of perceptions of events, commentators show a high degree of agreement in the structure and substance of their play-by-play reports. In quality, or entertainment value, there is perhaps more variation.

### 3.3 MANNER: HOW THINGS ARE SAID

Which brings us to *how* things are said, the details of form. It is here that we see most clearly the requirements of *quality* that are placed on commentaries. Descriptions must not only make sense to

the audience, they must entertain, being dramatic where drama is called for, humorous where humour is called for, brief where brevity is called for, and so on.

Various elements contribute to the quality of a play-by-play description, as we see in the following example from a game between Pakistan and Australia:

- (22) 3 *Mustaq - bowls again*  
 4 *to Taylor. -*  
 6 *It's overpitched*  
 8 *and TAYLOR HITS IT HARD! -*  
 9 *It's a GOOD hit. -*  
 19 *He COULD be CAUGHT! -*  
 19 *He SHOULD be CAUGHT! -*  
 21 *He's OUT! --*  
 20 *WELL taken by Waqar Younis.*

Here the sense of drama is created by several devices – features of grammar, wording, intonation, volume and tempo – which the commentator combines to build up tension as the events of play approach a climax, the catch by Waqar Younis.

### 3.4 MUSICAL PATTERNS

One set of dramatising devices has to do with the musical patterns of speech: pitch, intonation contours, volume and rhythm and tempo.

Kuiper and his co-authors point out that special prosodic modes are a characteristic of formulaic speech traditions, having the function (among others) of setting the discourse apart from ordinary speech. Within cricket commentaries we find well-marked shifts in prosodic patterns at certain points. For example, at a certain point in the description of each delivery – usually when the bowler begins his run or when he delivers the ball – the principal commentator usually moves into what I will call *dramatic mode* or *dramatic key*. The usual intonation pattern in dramatic mode is what Kuiper and Haggio (1984, 1985) and Kuiper and Austin (1990) term a 'drone' (indicated in example (22) and elsewhere by italics). The range of intonation in contour units is somewhat narrower than in 'normal' speech, with non-nuclear syllables pitched only slightly lower than syllables carrying nuclear stress, the pitch level is relatively high. The voice production has a quality that might be described as 'tense'.

In dramatic mode the rhythm and tempo are not quite those of ordinary speech. Each significant event, especially any of events 2-17, is usually sketched in a burst of two to five words, followed by a short pause. The pauses help to build up tension while we wait for the outcome of the previous event. The rate of articulation varies considerably between events. When describing the bowler's run-up and the act of delivery itself, an average (4.5 syllables per second) or slower than average rate of articulation is usual. But the immediate outcome of the delivery – the batsman's stroke, miss, etc.– is generally described at faster than average rate. Volume level usually starts off no louder than in other parts of the commentary but sharp increases in loudness often occur if the batsman plays a spectacular shot or if the batsman is out or if there is a chance he may be out.

If events that follow the batsman's stroke are seen as exciting the commentator will stay in dramatic mode. As a rule the transition back to normal tone contours and volume is made at either of

two points: (i) at the end of the play-by-play description, or (ii) at an earlier point when the outcome of the delivery is assured and is considered to be unexciting. In example (22) the excitement continues until the catch is made. In example (24), in section 4 below, it will be seen that the commentator stays in dramatic mode right into the score summary when describing a dramatic hit that brought the batsman close to his century and the team close to victory.

When the description of a particular event is longer than the usual two to five words, the rate of articulation in dramatic mode is likely to be faster than the normal speech rate.

### 3.5 GRAMMAR

Special effects are also created in play-by-play descriptions by grammatical devices. One of these is the use of tense. During this phase of a commentary, broadcasters mainly use simple present or present progressive to describe events – obviously so as to create the impression that they are relaying the events as they happen. During a longish play-by-play description commentators will often maintain this strategy even when they are actually speaking of events that happened two or three seconds earlier. One of the most common signals of a shift from play-by-play to colour commentary is a shift from present to past tense in descriptions of the action. In examples (3)-(5) above, we saw both a shift of tense and a shift of voice mode in the lines marked C.

'Telegraphic grammar' covers a variety of ways in which a commentator chooses to sketch an event with abbreviated clause structure. Recall the opening lines in commentary (3), repeated here for convenience:

- (3)            *McPhee, -- in, -*  
                  *to Waugh, --*  
                  *THIS time, -- short! -*  
                  *Waugh again just glides it*  
                  *down towards third man. -*

Example (3) might be paraphrased as (3'):

- (3')            McPhee runs in,  
                  and bowls to Waugh,  
                  THIS time the ball is short  
                  and Waugh again just glides it  
                  down towards third man.

Note the missing verb in the first, second and third lines of (3), the missing subject ('the ball') in the second line, the absence of an antecedent for 'it' (the ball) in line four), and the absence of conjunctions linking clauses. The resulting telegraphic style is effective – it is as if we are seeing flashes of the action, each too quick to be handled by a full clause.

Telegraphic grammar is an optional device, used frequently but not all the time. It is not necessary to abbreviate to the extent seen in (3) – in similar circumstances a commentator will often use full clauses.

## 3.6 WORDING

Let us turn now to wording. The most conspicuous elements of cricketing vocabulary are technical terms consisting of a single word or a compound. There are more than 2000 of these lexemes, and they can be found thickly distributed through any report on a cricket match, such as the passage from *Wisden* cited at the beginning of this paper. An example is *googly*, defined in Johnston and Webber (n.d.:60) as "a ball bowled out of the back of the right hand [i.e. with the back of the hand facing the batsman at the moment the ball is released] with what looks like a leg-break action, but which spins as an off-break when hitting the ground and so turns 'the wrong way' to what the batsman is expecting". This type of delivery is also called a *bosie* and a *wrong-un*.

Another major part of the vocabulary consists of standard clause-sized expressions which specify recurrent events in play. I will refer to these as *clause-sized formulas* or *formulaic constructions*. A few examples are given in Table 3.

TABLE 3: SOME CLAUSE-SIZED FORMULAS IN CRICKET COMMENTARY

(a)	[team] A <i>won the toss and elected to bat</i>
(b)	[team] A <i>won the toss and put</i> [team] B <i>in</i>
(c)	[batsman] A <i>be-TENSE on strike to</i> [bowler] B
(d)	[bowler] A <i>bowls to</i> [batsman] B
(e)	[the delivery] <i>be-TENSE (just) short of a length</i>
(f)	[batsman] A <i>comes/moves/is down the pitch</i>
(g)	[batsman] A <i>BE-tense beaten in the air</i>
(h)	[batsman] A <i>glides/chops/cuts (it) down to third man</i>
(i)	<i>The batsmen take a single</i>
(j)	[batsman] A <i>be-TENSE caught/dropped (by</i> [fielder] B)
(k)	[batsman] A <i>be-TENSE missed at</i> [position] x <i>(off</i> [bowler] B)
(l)	[fieldsman] A <i>put-TENSE down a regulation catch</i>
(m)	<i>Bowler A is on a hat trick</i>
(n)	<i>Umpire A calls "over!"</i>
(o)	<i>That's the end of the day's play</i>
(p)	<i>Rain has stopped play</i>

In Table 3 and elsewhere the following notational conventions are used in representing each formula: explanatory glosses, not part of the formula, are enclosed in [square brackets]; optional constituents are enclosed in (parentheses); all other constituents are obligatory; constituents whose lexical content is fixed are in *italics*; constituents whose lexical content is variable are in plain type; CAPITALS indicate a variable grammatical element; TENSE stands for variable marking of tense, aspect or mood; a slash between two constituents indicates a lexical choice.

In very few cases is the wording of clause-length formulas rigidly fixed. Most clause-sized formulas consist of a grammatical construction whose word content is partly fixed and partly variable. That is to say, each formula is a schema for generating a family of expressions or word-strings that are functionally equivalent, in terms of the job they do in the subject matter code. While the sketches of each incident are stylised and tend to be similar, there is room for variation and novelty. For example, a common routine for describing the bowler's run to the wicket is:

- (23) A  $\left\{ \begin{array}{l} \text{VERB-s} \\ \text{comes VERB-ing} \end{array} \right\}$  in (to bowl to B)

Here *VERB* may be any one of a number of verbs: *move*, *run*, *come*, *stride*, or a verb phrase beginning with *comes*, for example, *comes running in*, *comes steaming in*, etc. Several examples of productive formulas for specifying scoring details can be seen in Table 5 below.

There are at least several hundred clause-length formulas that are part of the distinctive vocabulary of cricket. A full account of the cricket commentator's vocabulary would include a list of all of these together with a statement of the discourse contexts in which each formula is properly used and a detailed discussion of the grammar of each.

Some formulas in players' discourse are not descriptive but performative. That is, they are uttered as a part of play, for example, in order to influence or decide an outcome. An example is the phrase *How was that?*, a conventional way of appealing to the umpire for a decision against a batsman. The phrase is properly used by any member of the fielding team when something has happened that might warrant such an appeal, for example, when the batsman has been hit on the pad by a straight delivery, or when he may have deflected the ball with his bat to a fielder. In this context the umpire should respond with another speech formula or a conventional gesture. The phrase *Not out!*, serves as a ruling against the appeal; raising the index finger gives the batsman out.

The formulas used by commentators are descriptive. But 'descriptive' is perhaps not the best word. As already noted in 1.1, all descriptions are necessarily sketches, shorthand ways of getting across an idea. The descriptions in play-by-play reports, in particular, are packaged for a fast ride. The commentator is under considerable time pressure and can mention just a few details. Formulaic expressions are well suited to the needs of play-by-play commentary because they provide the broadcaster and audience with ready-made word-pictures.

For every recurrent event-type there is often a choice available between two or more formally quite different formulas that are functional equivalents, such as: '*Umpire X raises his finger*', '*Umpire X gives him out*', '*Umpire X says "That's out!"*' '*He's been given out!*', etc.

Words and phrases that elaborate on events instead of describing them as starkly as possible in strictly technical or 'objective' terms – for example, '*He SMASHES it*', '*BRILLIANTLY stopped by Jones!*', '*What a fantastic catch!*' – have *evaluative* force, in the sense of Labov (1973). Evaluative work is done not only by adjectives, adverbs and the like, which modify the bare-bones description provided by the verb and nouns but also by clauses which state not what did happen or is happening but what *might* happen, what *almost* happened, what *will* happen, etc. The most striking bit of evaluative work in example (22) is done by the three clauses '*It's a GOOD hit!* - *He COULD be CAUGHT!*-, *He SHOULD be CAUGHT!*' which create suspense while we wait for the outcome of the lofted stroke. The final clause evaluates the catch: '*WELL taken by Waqar Younis*'.

In the vocabulary of cricket generally and in the wording of commentaries we can see that the main point of reference, the focal point of play, is the batsman, or the batsman's wicket. For example, commentators often speak of the bowler '*coming* (rarely *going*) *in to bowl*'. The three stumps and the positions on the field are named with reference to the batsman's orientation or his wicket: *leg stump*, *off stump*, *leg side*, *off side*, *square leg*, *mid-on*, *mid-off*, *point*, *cover*, etc. So too, not surprisingly, are some of the terms for movement, direction and length of the ball in a delivery: *leg-*



*break, off-break, inswinger, outswinger, bodyline, full toss, etc.*, and the terms for many of the batsman's strokes.

#### 4. SCORE SUMMARIES

While play is in progress it is the principal commentator's job to keep the listener informed about statistical details of the game by means of regular reports, which I will refer to as *score summaries* or *score reports*. A full account of the structure and use of score summaries would require a paper by itself; I will offer just a few notes here.

A score summary may report any one of upwards of twenty categories of statistical information, such as the scores of the two batsmen who are currently batting, the total score of the batting team, the number of wickets fallen (i.e. of batsmen out) so far during the innings, the runs scored by each batsman already out, the team totals at the fall of each wicket, the number of runs accumulated during the current or past batting partnerships, scoring rates (in terms of number of runs per minute or per deliveries received), the over rate of the fielding team, the number of runs conceded by a bowler in an over, the number of wickets taken and runs conceded by a bowler in an innings, the batting team's targets to win (in terms of runs needed or overs or deliveries remaining), the number of runs needed to avoid the follow on (when a team is asked to take its second innings immediately after its first because its total is far in arrears of the team that batted first in the game), and so on.

Score reports occur in fixed contexts relative both to the state of the game (external or game contexts) and to the content of the discourse as shown in Table 4.

TABLE 4: GAME CONTEXTS IN WHICH SCORE SUMMARIES ARE GIVEN

(a)	after each delivery has been described	(optional)
(b)	after each ball that has conceded a run	(obligatory)
(c)	at the end of each over	(optional)
(d)	at the fall of a wicket	(obligatory)
(e)	during drinks (one hour into each session)	(obligatory)
(f)	at the end of each session of play	(obligatory)

Score reports show all the diagnostic features of oral formulaic speech. Their discourse structure is extremely well defined, in terms of the subject matter mentioned and the ordering of the information. There is a formula (more often two or three) to express each type of statistical information. Table 5 lists a small selection of individual scoring formulas, with the letters x, y, etc. representing variable scores (number of runs, wickets, etc.). Words in square brackets are explanatory, and are not part of the spoken formulas.

TABLE 5: SOME SCORING FORMULAS

(a)	[batsman] A is ( <i>not out</i> ) x [runs]
(b.i)	[team] A is x ( <i>wickets</i> ) ( <i>down</i> ) for y [runs] [preferred by Australian commentators]
(b.ii)	[team] A is x [runs] for ( <i>the loss of</i> ) y ( <i>wickets</i> ) [preferred by U.K. commentators]
(c)	[team] A is x [ <i>wickets</i> ] ( <i>down</i> ) for y [runs] <i>from/off</i> y <i>overs/balls</i> [(c) is used mainly of one-day matches]



- (d) *After x overs the score/total is y* [runs]  
 (e) *(There were) x runs (taken) off/from the over*  
 (f) [bowler] *A conceded x (runs) off/from the over*  
 (g) [batsman] *A was (out) caught* [fielder] *B, bowled* [bowler] *C for x* [runs]  
 (h) [team] *A was / were all out / dismissed for x* [runs]  
 (i) [bowler] *A take-TENSE x (wickets) (for y* [runs])  
 (j) [bowler] *A (bowled-TENSE) u overs, (including/with) v maidens, and take-TENSE w (wickets) for x (runs)*  
 (k) *The partnership (for the nth wicket) be-TENSE worth x runs*  
 (l) [batsmen] *A and B (have) added / put on x runs (for the nth wicket)*

For most kinds of scoring information the commentator can choose between a formula with telegraphic grammar and one with 'full' grammar, for example, 'England has scored 98 for the loss of two wickets' versus 'England 98 for two', 'Boycott was out caught Chappell off the bowling of Walker for 41' versus 'Boycott caught Chappell bowled Walker, 41'.

Most score reports have special prosodics. In contrast to play-by-play descriptions, they are usually spoken slowly, with exaggerated clarity, both stressed and unstressed syllables being given more prominence than usual with each digit carrying a tone peak. This is especially true of summaries which combine several scoring formulas or which consist of a single complex formula. The total effect is of rather singsong speech, which signals to the listener: 'I am saying all this carefully because you may want to note every detail'.

The most frequent score reports are updates of the batsmen's and batting side's totals. The normal place for giving these is at the end of each play-by-play description, either before or after the colour commentary – most often immediately before or after the first recapitulation of events in the last delivery. In examples (3)-(5), given earlier, the summaries followed the recapitulation. However, in example (24) below the summary was given first – one reason for this was that the batsman was close to completing his century, and his progress towards this target had become a prime focus of attention for the commentators.

- (24) P 1,3 *Border bowls to him again, -*  
 8 *and he LOFTS that away! -*  
 9 *He's punched that HARD and HIGH, -*  
 11 *into the crowd*  
 18 *for six. -*  
 S *NINety NINE for John Wright. -*  
*One for one fifty NINE New Zealand, -*  
 C *He's enjoying Allan Border's bowling now. -*  
*That was a suPERB blow! -*  
*That REALLY WAS!*

Notice that the dramatic mode is continued from play-by-play description into the score summary in accord with the crowd's excitement as a spectacular stroke by Wright moves him to within one run of his century and the New Zealand team within a few runs of victory.

## 5. COLOUR COMMENTARY

The dramatics of radio cricket commentaries are not nearly as sustained as those of commentaries on fast, continuous action sports like basketball, soccer, ice-hockey and horse-racing. To the uninitiated onlooker a game of cricket seems to be made up of brief bursts of action punctuated by long periods in which nothing happens. Furthermore, all of the bursts of action begin with the same routine: the bowling of the ball to the batsman.

For the informed spectator, however, the cyclic nature of play and the intervals between deliveries and overs are central to the game's aesthetic appeal. The intervals provide time for the build-up of tension and spaces for reflection and conversation on tactics, techniques and individual styles of play. It is unlikely that people would listen to commentaries for hour after hour if they consisted wholly or predominantly of play-by-play descriptions and score summaries. Much of the charm of cricket broadcasts resides in the colour commentary, where the two commentators may talk over the immediately preceding events or discuss any one of 50 other topics: the strategic options of a captain, batsman or bowler, the technical skills or limitations of a particular player, the state of the pitch, the quality of the umpiring or the conduct of the crowd. If nothing at all is happening on the field – as when rain stops play – the commentators can reminisce about the heroes and heroic deeds of the past or speculate on the thoughts of the seagulls watching proceedings from the outfield.

The change from play-by-play to colour commentary is, as noted in 3.4, signalled by a number of features. The most frequent markers are shifts in the commentator's voice quality and intonation from dramatic mode to a more relaxed style, change from present to past tense in verbs describing the events of play, a shift from highly formulaic, telegraphic descriptions of events to longer, more elaborate descriptions. Some of these stylistic shifts can be seen in excerpts (3)-(5) and are clearly illustrated in the following piece of commentary from a Test match between Australia and New Zealand. The play-by-play commentary, in full dramatic mode, is followed by a brief score summary, only slightly less dramatised. There is a marked change when the commentators begin to review Jones's stroke – the volume drops, the pitch range widens and the grammatical constructions become more varied, with questions, interpolations and past tense verbs.

- (25) P 4,7 *ANDREW JONES GOES AFTER*  
 3 *DEAN JONES!* -  
 8 *HE SMASHES THAT*  
 12 *OUT THROUGH MID-WICKET*  
 17 *FOR FOUR!* -  
 19 *Carl Rackemann gives up the chase very early in the piece.*  
 S *One for one hundred and sixty three!* -  
*Andrew Jones is THIRTY THREE,*  
*John Wright is NINETY NINE.* -  
 C1 He really hit that one HARD-  
 He swung--maybe slightly across the line was it? -  
 But whatever it was,  
 it was an enormously aGGRESSive shot down on the knee  
 and he BLASTed that out to the mid-wicket boundary. -  
 It was just in front of square leg.  
 [The colour commentator chimes in here:]

- C2 Wasn't such a bad ball from Dean Jones, really. -  
 It was middle and leg.  
 But Andrew Jones picked the right length  
 and really hit it solidly.

For other examples, see (3)-(5) and (24).

Neither the topics nor the wording of colour commentaries as a whole are as predictable overall as play-by-play reports or score summaries. However, the limited analysis I have done on colour commentaries suggests that they can be broken down into a number of constituents or subtypes, each of which has a fairly well-defined internal structure. Among these highly-structured elements are *announcements of the toss for innings, descriptions of the state of the pitch, field setting descriptions, mention of personal milestones, moral evaluations (on sportsmanship, courage, etc.), recapitulation of events in a unit of play and announcements of a change in commentator.*

The discourse rules of colour commentary are complicated by the fact that two speakers must take turns. The convention is that the principal commentator provides the play-by-play description and the score update and also has first crack at recapitulating the events of the delivery. The colour commentator may then chime in with additional observations and from then on the commentary may proceed like any conversation. But we must leave closer analysis of colour commentaries for another day. It is time to move to broader questions.

## 6. EXPLAINING THE STRUCTURE OF RADIO CRICKET COMMENTARIES

We have examined three components of radio cricket commentaries. Two of these, play-by-play descriptions and score summaries, exhibit all of the features characteristic of formulaic speech genres:<sup>13</sup> (a) strict discourse structure rules specifying the topics that may be mentioned and their order, (b) a high concentration of speech formulas for expressing the things that may be said about each topic, (c) special grammar, (d) special musical patterns and (e) fewer than average unplanned pauses in mid-construction. The third component, colour commentary, is less tightly structured but appears to be analysable into a number of components some of which are highly formulaic.

I turn now to the *explanation* of the structure of cricket commentaries. Why do the commentators follow extremely strict discourse structure rules? Why do they make such extensive use of formulas? Why are special prosodics used in certain parts of the commentary? These questions have already been touched on at various points in the preceding discussion, where we have discussed the functions of particular usages, but some general remarks are in order.

The well-defined discourse structure of play-by-play descriptions has surely arisen because play in cricket is cyclic. The basic units of play are the delivery and the over, consisting of six deliveries. A day's play contains around 90-120 overs or 540-720 deliveries, and as similar things happen in many deliveries, the constituent subject matters of play-by-play reports recur throughout any match.

In keeping with earlier work on oral formulaic discourse,<sup>14</sup> I suggest that commentators' reliance on speech formulas to describe specific events in play-by-play reports can be explained in terms of

<sup>13</sup>It is wrong to regard this kind of analysis as being concerned simply with linguistic 'performance', i.e. with the production of discourse, rather than with an underlying system. The subject matter is conventions, principles, rules.

<sup>14</sup>See, for example, Becker (1975), Bolinger (1976), Chafe (1977, 1979, 1980), Kuiper and Haggio (1984), Kuiper and Austin (1990), Pawley and Syder (1983).

three main factors: (i) the requirement of economy or speed, (ii) the requirement of quality, (iii) the nature of human mental processing capacities.

The processing load on the commentator is heaviest in play-by-play reports. The task is to instantly comprehend the action and with minimal delay to transform one's perceptions into an accurate, easily understood and entertaining description. During descriptions of particular events the chunks of information that are manipulated in each focus of consciousness presumably include visual ones – catching the flight of the ball, the batsman's movements, etc. – as well as the cognitive ones of working out what is significant and putting the concepts into words.

The short term memory – roughly what we can hold in our minds in a single focus of consciousness – is very limited; experimental research suggests that it is roughly seven chunks of unrelated information. A person's long-term memory capacity, on the other hand, is virtually boundless. Therefore it is an advantage, when called on to produce or to understand rapid descriptions of a series of events, to have stored in the long term memory standard knowledge schemata, binding together event contexts, conceptual structures and descriptions into a single Gestalt or chunk. Instead of having to break his observations of events down into constituent concepts and to retrieve a word appropriate to express each concept, the commentator can match the observations with schemata and draw on a suitable descriptive formula. The listener, hearing a shorthand verbal sketch of an event, can reconstruct many of the details from his backing knowledge. The wording and the phonology of formulas are not only familiar, they have been streamlined for economy of articulation.<sup>15</sup>

When announcing statistical details brevity and accuracy, rather than drama, are the key requirements and this weighting is reflected in the extremely stereotyped but efficient form of score reports and in their generally deliberate articulation. In colour commentary the main concerns are different – to elaborate and entertain – and these emphases are reflected in the more diverse discourse structure, formulas, grammar and prosodics found in this domain.

## 7. ON APPLYING THE GENERATIVE FORMULAIC MODEL TO 'ORDINARY' LANGUAGE

We have sketched a framework for describing a particular subject matter code, namely, radio play-by-play commentaries on cricket. The framework is intended to be generative, stating explicitly the kinds of knowledge a person must have to be competent in such a genre. The user must have, first, a rich encyclopaedic or backing knowledge of the cricket culture and must know (relevant parts of) the conventional grammar and lexicon of the language. But in addition, he must know the conventions regarding:

(a) The agenda of discourse: that is, the topics it is appropriate to talk about and their structural relationships. This information may be described by a set of discourse structure rules specifying topics and contextual constraints, for example, on sequencing and mutual exclusiveness;

(b) The specific things that are sayable. The claim is that component (b) can largely be specified in terms of speech formulas indexed to discourse structure contexts. Each speech formula is a standard recipe for binding linguistic form, meaning and purpose. As noted in section 3.6, a formula may be

<sup>15</sup>The special prosodic features which occur in play-by-play commentaries are best considered not as part of individual formulas but as variables governed by discourse purpose and context. Their main function is not to ease the load on the short term memory but to colour the discourse, to make it hold the listener's attention.

lexically fixed or variable. A variable formula is in fact a mini generative grammar, in which certain conceptual and formal elements in the construction may be varied by rule or list, although its broad semantic and grammatical structure and its discourse function remain the same.

Other studies have shown that the main features of this framework are applicable to a variety of subject matter codes that are often termed 'oral formulaic genres'. The question arises: can the descriptive model that has been devised for analysing formulaic speech genres be applied with profit to the analysis of 'ordinary language'?<sup>16</sup>

The first thing to be said is that it is misleading to formulate the question in this way. For one thing, 'ordinary language' is not a well-defined category, something that can be contrasted with 'formulaic speech'. What we take to be everyday language is a very mixed bag. It is safe to say, however, that the most common and universal mode of language use is conversation. And while most conversation is not purely formulaic, the very large body of research on conversation carried out over the last 25 years has shown that everyday talk is studded with standardised interchanges. When people give sermons or tell jokes, when they buy and sell, or seek to control, persuade, apologise, forgive, blame or praise, congratulate, commiserate, curse or bless, ask for a favour or refuse a request, their talk nearly always follows fairly strict rules of discourse structure and is heavily spiced with formulas. All hardly surprising, given that everyday life consists largely of familiar, recurrent situations. It is essential that people have tried and true ways of categorising and talking about such recurrent situations.

In cricket broadcasts, conversational speech is represented chiefly by the colour commentary interludes between play-by-play descriptions. We noted in section 5 that while such interludes, taken as a whole, are often less tightly structured than play-by-play commentaries, they appear to be largely made up of constituents that are highly formulaic. It would, however, take a separate paper to substantiate this claim in detail.

It happens that the term 'formula' has a bad press. It is often thought to stand for bits of language that are completely fixed in structure and therefore uninteresting except to phrase collectors.<sup>17</sup> But this judgment is by no means accurate or fair. For one thing, creative behaviour of any kind itself depends on the existence of norms. It is precisely our knowledge of the standard schemas that allows us to appreciate the departures from these. People like novelty but they prefer it a little bit at a time, in the form of variations on familiar themes. The expert performer in an oral formulaic tradition knows how to exploit this fact in order to surprise and entertain an audience.

This is as true of good cricket commentators as it is of good talkers in other fields. Let me give just one example. The commentator was the peerless John Arlott of the rich Hampshire voice, the game was a Test match between England and South Africa in 1949. When Tufty Mann, the South African bowler, bowled out George Mann, the English captain, Arlott exclaimed: "*He's bowled the England captain! Such is Mann's inhumanity to Mann!*"

<sup>16</sup>The question of how far the formulaic treatment of the composition and performance of oral poetry, as described, for example by Lord (1960), can be applied to everyday speech has been raised by a number of writers, including Kiparsky (1976), Kuiper and Hago (1984) and Pawley and Syder (1983).

<sup>17</sup>For example, Quirk et al.'s (1972) *A grammar of contemporary English* devotes just two pages out of 1100 to formulas, placing them in a section headed 'a museum of oddments'.



## 8. CONCLUSIONS

While I have written about a game my purpose in this discussion has been deadly serious. I hope, first, to have defined and illustrated the notion of a subject matter code. Second, I hope to have given evidence that a rigorous analysis of at least some subject matter codes is attainable – for example, that it may be possible to identify the significant elements in such commentaries and to arrive at an explicit, exhaustive statement of their possible combinations. Third, I hope to have convinced the reader that to do so we need an analytic framework that is considerably richer than the conventional grammar-lexicon model.<sup>18</sup>

Grace (1987:104) writes that “the extent of our *understanding* of a particular area of subject matter is equal to a kind of sum of our ways of talking about that thing... Or to put it differently..it is equal to a kind of sum of the number of its factors we have been able to bring into perspective”. It is precisely our ways of talking about language, he argues, that creates our view, as linguists, of what language is.

Useful though it is for certain purposes, the tradition of speaking of languages as if these consist primarily of a grammar, incorporating a lexicon which is an appendage of irregularities, is a Procrustean bed that leads us to disregard some of their most important and puzzling elements. Grammatical systems are empty codes, in the sense that they tell us little about people’s knowledge of and beliefs about the world or about how people actually talk about things.

I believe the focus on subject matter codes brings into perspective certain neglected ingredients of language. More clearly than morphology and syntax, discourse structure, speech formulas and vocabulary stand at the interface between language and culture, because they reflect people’s ways of talking about particular subject matters. We are dealing here with domains of codified knowledge and verbal art that are often the product of centuries of accumulated experience. It happens that we live in an age in which many longstanding language-culture systems are in danger of disappearing virtually without trace. I suggest that, apart from their theoretical interest in such systems, linguists have a wider obligation to record the full richness and diversity of Mankind’s linguistic creations. To do so, however, we need not only to apply our traditional analytic techniques but to extend them.

## BIBLIOGRAPHY

- BECKER, Joseph D., 1975, The phrasal lexicon. In Schank and Nash-Webber, eds 1975:60-63.
- BOLINGER, Dwight, 1976, Meaning and memory. *Forum Linguisticum* 1/1:1-14.
- CHAFE, Wallace L., 1977, The recall and verbalization of past experience. In Roger W. Cole, ed. *Current issues in linguistic theory*, 215-246. Bloomington: Indiana UP.
- 1979, The flow of thought and the flow of language. In T. Givon, ed. *Syntax and semantics*, vol.12: *Discourse and syntax*, 159-181. New York: Academic Press.
- 1980, The deployment of consciousness in the production of a narrative. In W.L. Chafe, ed. *The pear stories: cognitive, cultural, and linguistic aspects of narrative production*, 9-50. Norwood, NJ: Ablex.
- FREWIN, Leslie, ed., 1962, *The boundary book: a Lord’s Taverners’ miscellany of cricket*. London: Macdonald.
- GRACE, George W., 1981, *An essay on language*. Columbia, S.C.: Hornbeam Press.
- 1987, *The linguistic construction of reality*. New York/London: Croom Helm.
- JOHNSTON, Brian and Roy WEBBER, n.d., *Armchair cricket*. London: British Broadcasting Corporation.

<sup>18</sup> I borrow the term ‘grammar-lexicon’ model from Grace (1981).



- KIPARSKY, Paul**, 1976, Oral poetry: some linguistic and typological considerations. In B.A. Stolz and R.S. Shannon, eds *Oral literature and the formula*, 73-106. Ann Arbor: Center for the Coordination of Ancient and Modern Studies, University of Michigan.
- KUIPER, Koenraad and Paddy AUSTIN**, 1990, They're off and racing now: the speech of the New Zealand race caller. In Allan Bell and Janet Holmes, eds *New Zealand ways of speaking English*, 195-220. Clevedon: Multilingual Matters.
- KUIPER, Koenraad and Douglas HAGGO**, 1984, Livestock auctions, oral poetry, and ordinary language. *Language in Society* 13:205-234.
- 1985, On the nature of ice hockey commentaries. In Reginald Berry and James Acheson, eds *Regionalism and national identity: multi-disciplinary essays on Canada, Australia, and New Zealand*, 167-175. Christchurch: Association for Canadian Studies in Australia and New Zealand.
- KUIPER, Koenraad and Frederick TILLIS**, 1985, The chant of the tobacco auctioneer. *American Speech* 60:141-48.
- LABOV, William**, 1972, *Language in the Inner City: studies in the Black English vernacular*. Philadelphia: University of Pennsylvania Press.
- LORD, Albert**, 1960, *The singer of tales*. Cambridge, Mass.: Harvard University Press.
- PAWLEY, Andrew**, 1985, Speech formulas and linguistic competence. *Linguas Modernas* 12:80-104.
- PAWLEY, Andrew and Frances H. SYDER**, 1983, Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In Jack C. Richards and Richard W. Schmidt, eds *Language and communication*, 191-225. London: Longman.
- QUIRK, Randolph, Sidney GREENBAUM, Geoffrey LEECH and Jan SVARTIK**, 1972, *A grammar of contemporary English*. London: Longman.
- SCHANK, Roger C. and B.L. NASH-WEBBER**, eds, 1975, *Theoretical issues in natural language processing*. Cambridge, Mass.: Bolt Beranek and Newman.
- SNAGGE, John**, 1962, The voice of cricket. In Frewin, ed. 1962:131-132.
- WEST, Peter**, 1962, Room with a view. In Frewin, ed. 1962:132-144.

TONGAN SPEECH LEVELS: PRACTICE AND TALK ABOUT PRACTICE  
IN THE CULTURAL CONSTRUCTION OF SOCIAL HIERARCHY

SUSAN U. PHILIPS

1. INTRODUCTION<sup>1</sup>

The purpose of this paper is to consider the relation between Tongan conceptualisations of speech levels and Tongan use of speech levels. I will argue that while both are concerned with distinctions among speech levels, these two cultural constructions of social hierarchy differ in important and interesting ways. This paper should be considered a preliminary introduction to the topic, with additional and more detailed analysis to follow in later writing.

While it has been written about Tongans that they use different terms to refer to the same actions of Commoners, Nobles and the King, in practice words for both kingly and noble actions are extended to other high status social categories. But it is primarily noble terms that are widely used. They play an important role in respectful public speech in general, and in performed verbal art in particular. Thus while statements about these three separate vocabularies suggest they are analogously distinct from one another, the kingly and noble terms are very different from one another in patterns of use, and play rather different roles in the Tongan cultural construction of social hierarchy. Moreover, discussion of Tongan speech levels (and of speech levels in other societies) usually suggests that they can be understood without reference to other sources of variation in language use, so that they are kept separate conceptually from other kinds of analysis of language use by Tongan and non-Tongan analysts alike. But it is clear that in Tonga the tradition of performed verbal art is in practice very much intertwined with the use of Tongan speech levels.

---

<sup>1</sup>The research on which this paper is based was supported by the Wenner-Gren Foundation in 1985 and by the NSF Linguistics Program in 1987 and 1988. Others besides myself who have contributed to data collection and processing relevant for this paper include Ernest Olson, Charles Stevens, Keiko Matsuki, Liliani Maile, Toti Maile and other Tongan research assistants. I also wish to express my gratitude to the Government of Tonga for allowing this research to be carried out, to the family of Vili Salakielu for their hospitality and support, and to Eric Shumway and Tongan consultants Futa Helu and Manu Faupula for their instruction in Tongan language and culture. None of these people can, however, be held responsible for any problems with this discussion.

By speech levels, I refer to the aforementioned three-way distinction among vocabulary items used to and about Commoners, Nobles and the King. These have been presented in very similar formats, partially reproduced in Table 1, in columned word lists in Gifford (1929), Shumway (1971) and *Ko e Kalama 'i he Lea Faka-Tonga* (n.d.) that show the parallel equivalents of commoner, noble and kingly terms, and in the case of Gifford and Shumway, their English equivalents. What I am calling speech levels here have also been referred to as respect or honorific forms of language use. I choose the phrase 'speech levels' to convey their very paralleled and hierarchical conceptualisation, analogous to Javanese speech levels (e.g. Errington 1985), even though Tongan levels are characterised primarily in lexical terms, without the phonological and morphosyntactic dimensions of complexity readily recognised in the description of Javanese speech levels.

TABLE 1: EXAMPLES FROM PUBLISHED TONGAN SPEECH LEVEL WORD LISTS

1a: From E.W. Gifford (1929) *Tongan society*: Rank and Language

1. MUOMUA (LEADING CHIEFS)	2. LOTOLOTO (MIDDLE CHIEFS)	3. KAKAI (PEOPLE)	ENGLISH
fakamomoko	fakamomofi	fakamafana	To warm one's self
fakamonuka	polai	tafai	To lance (cut)
feitaumafa	halofia	fiekai	Hungry
fekita	uma	uma	To kiss
hala	pekia	mate	Dead
hoihoifua	katakata	kata	To laugh
hoko	vala	fatei	Skirt-like garment (sarong)
houhau	tupotamaki	ita	Angry

1b: From E. Shumway (1971) *Intensive course in Tongan*: Respect Terminology for Nobility and Royalty

English	Kakai	Hou'eiki	Tu'i
angry	'ita	tuputāmaki	houhau
bathe	kaukau	tākele	fakamālū
cemetery	fa'itoka	mala'e	mo'unga
children	fānau	fānau	fale'alo
clothes	kofu	kofu	fakama'u
come	ha'u	me'a	hā'ele
converse	fetalanoa'aki	feme'a'aki	fefolofolai
desire	loto	loto	finangalo
die	mate	pekia	hala
eat	kai	'ilo	taumafa

*Table continued on next page*

1c: From *Ko e Kalama*: Vocabulary for the Three Classes of People in the Tongan Society, Commoners, Nobles and the Royalty

COMMONERS (Kakai)	NOBLES (Hou'eiki)	KING (Tu'i)
manako	manako	mokoi
fata	vaka	kauala
navu	pepena	palai
mohevale	misi	lika
ngangau	mamafa	ngalulu
mate-'i-tahi	mole-'i-tahi	takavaha
fulufulu	halakava	hoha'a
hinehina	hinā	sialea

Tongan speech levels are recognised by Tongans and Western European linguists alike (e.g. Milner 1961; Blixen 1966) to be similar to and related to Samoan respect vocabulary distinctions between commoner and chief, and partially also between *tulafale* (talking chiefs) and highest chiefs (Duranti 1987). They are also recognised in the linguistic literature (Tregear 1895; Milner 1961) to be instances of a phenomenon widespread in Austronesian languages of the linguistic marking of relative status of speakers or relative importance of speech and speaking context, although the linguistic realisations of such distinctions and their local cultural conceptualisations vary greatly (see, for example, Watson-Gegeo and Gegeo 1986) and there has been considerable disagreement about how such distinctions are historically related.

Silverstein (1979, 1981) has used the term 'metapragmatics' to refer to the analysis provided by speakers of a language of the ways in which their language is used. He argues that speakers are able to identify certain aspects of language structure and associate them with functionally differentiated patterns of use more readily than others. In other words, some aspects of language form are more salient and subject to metapragmatic analysis than others. This point can be seen as commentary on and qualification of Boas's (1911) view that language is generally less subject to "secondary rationalizations" than other forms of culture, and hence a purer and better source of data for establishing prehistoric relations among social groups. More particularly Silverstein suggests that segmentable phenomena such as lexical items and isolatable morphemes are more accessible to speakers than non-segmentable phenomena in language such as phonetic or phonological variation and syntax.

Shibamoto (1987) has addressed this same issue in her analysis of Japanese concepts of the ways in which men's and women's language use differs in relation to behaviour. Segmentable elements, such as the greater use of the  $\sigma$ - prefix on nouns by females and gender-differentiated use of pronouns and sentence-final particles are salient in Japanese gender metapragmatics. In contrast, non-segmentable syntactic variation in word order, while documented by Shibamoto in analysis of naturally occurring speech, is not part of Japanese gender ideology. Shibamoto has also been able to show that those aspects of linguistic behaviour do not vary across contexts in the same ways, by comparing conversational data with the Japanese equivalent to television soap operas. Soap operas are recognised to exaggerate or intensify feminine and masculine characteristics of speech in everyday life. Shibamoto found that the features of speech that were part of Japanese metapragmatics of gender differences in language use were in fact exaggerated in the television programs compared to conversational data. But those aspects of linguistic behaviour which she found to behaviourally

correlate with gender that were not part of this gender ideology were no different on television than in conversation.

Analogously, Errington (1985) has argued that the segmentable aspects of speech level usage most elaborated in Javanese metapragmatics, namely the usages of second person pronouns, are undergoing change over time in a different way than those non-segmentable aspects of speech level usage that are less salient to speakers and less elaborated metapragmatically.

I use these studies here first to make the point that Tongan metapragmatic analysis of speech levels is consistent with Silverstein's suggestion that local metapragmatics are more likely to focus on segmentable phenomena, in that characterisation of speech levels is overwhelmingly in terms of analogous lexical items, rather than syntactic or phonological phenomena. And like Shibamoto and Errington, I wish to distinguish between theories of use and use or practice itself. However, unlike them, I will stress that there is also theory *in* use, but of a different nature or order than theories of use.

But there is an additional concern to be addressed here initially in this metapragmatic turn of analysis, and that is that it raises the possibility of a distinction between local members' metapragmatic analysis and the analysis of the linguist. In the three sources on Tongan speech levels that I have cited so far (Gifford, Shumway and *Ko e Kalama*; cf. Table 1), which I view as examples of metapragmatic analyses, the basic tripartite distinctions among levels are basically the same, although variation among them will be considered later. This is particularly true of the two more recent and currently-in-use presentations of Shumway and *Ko e Kalama*, which use the English words Commoner, Noble and King as the social identities with which each parallel lexical list is associated. Here it appears that Tongans and Western scholars of their language embrace essentially the same view of these speech levels.

How does such agreement come about? It is not the purpose of this paper to dwell on this issue, nor am I in a position to trace the development of such agreement. But I believe that for Tonga, and I suspect that for other parts of the world, local interpretations of local cultures are often quite directly taken up by Western European scholars and presented as their own, without any distinction being made between the two, or 'voices' being carefully sorted out. Certainly much of Gifford's analysis of kinship organisation, and that of more recent scholars, sounds very much like the kinds of things Tongans say today about the mother's side and the father's side. And Marcus's (1980) explanation of the new commoner ways of achieving status arising alongside of and in some ways partially replacing traditional avenues to status is also commonly spoken of in Tonga today. It can be argued, of course, and many contemporary anthropologists do so argue, that contemporary local cultural analyses are the result of a dialectic over time between coloniser and colonised. Thus one might envision Gifford's informants having given him a certain account, which he then modified, with that account then influencing later accounts up to the present. Here I simply want to suggest more Tongan agency in the creation of these accounts, both in the past and in the present, than is typically recognised. Today *Ko e Kalama* is the text book used by Tongan high school students for instruction in the proper use of Tongan speech levels, use which they are examined on in the government-mandated exams taken by Tongan high school students that then determine whether they will have access to particular kinds of government jobs, which are in turn the main source of cash income in this tiny non-industrialised nation. *Ko e Kalama* and its use, then, are government-supported representations of Tongan speech levels, and most of the Tongans with whom I discussed speech levels referred to this high school training as a foundational source of their knowledge of such terminology.

## 2. TONGAN TALK ABOUT PRACTICE

In Tonga today, and in the literature on Tonga, metapragmatic discussions of language use focus on hierarchical aspects of Tongan society. In a general way, the linguistic expression of social hierarchy is conceptualised in terms of the use of a variety of verbal strategies for raising others and lowering oneself. For me such strategies are epitomised in what in Western terms is the host-guest relationship, but here it may also be appropriate to think in terms of an insider-outsider relationship. Insiders to social units of varying sizes and natures, but often most publically and visibly the village, relate as hosts to outsider guests, raising their guests and lowering themselves.

Analytically, in Tongan talk about talk, there are at least two manifestations of hierarchically organised language use, in each of which lexical characterisations figure prominently. First, Tongans distinguish between words to and about Commoners, Nobles and the King, to be elaborated further on. Churchward (1953) elaborates on the three-way distinction by positing five levels: 'vulgar', 'everyday', 'polite', 'honorific' and 'regal'. More recently, Tali'ai (1989) has described Tongan speech levels in a four-way system, distinguishing among *lea 'oe tu'i* 'language of the king', *lea 'oe hou'eiki* 'language of the chiefs', *lea 'a e tu'a* 'language of the commoners', and *lea 'oe tatau* 'language of the equal'. In this framework, many of the *tu'a* words would be characterised as vulgar in Churchward's terms, while *tatau* words are those associated with Commoners, who are referred to as *kakai* 'people' in Tongan in other analyses (cf. Table 1), rather than as *tu'a*. *Tu'a* actually comes closer to English 'commoner' in that it refers to untitled persons and definitely carries the implication of lesser status than chiefs, and this is probably the reason it is not usually used. It should also be noted that both Collocott (1925) and Blixen (1966) have a predominantly dualistic conceptualisation of Tongan speech levels, possibly influenced by the predominantly dualistic conceptualisation of the Samoans with regard to their respect vocabulary.

These interpretations suggest some variability and instability in the way speech levels are conceptualised. In my experience it is common for Tongans to distinguish between everyday and polite usage, and to identify particular words as impolite or vulgar. But the three-way differentiation I discuss here is still by far the more common and salient way of identifying status-differentiated speech levels.

The second major distinction relevant to the cultural construction of social hierarchy in Tongan talk about language use is that between everyday language use and what English speakers refer to as 'public speech', or the speech associated with what is rightly called the Tongan literary tradition, epitomised in the works of *punake*, composers of poetry and song-dances that entail the combined or separate abilities of lyricist, composer and choreographer (Shumway 1977; Kaepler 1976), and also in the talk of *matāpule*, the titled talking chiefs of Tonga. This literary and rhetorical tradition would in contemporary sociolinguistic terms be considered speech that is performed (Bauman 1975). It is verbal art, subject to audience evaluation and appreciation, the greatest gift of host to guest. People are moved by its beauty, and held spellbound by its wit, intellect, and resourceful display of knowledge of Tongan myth and history.

Like the first metapragmatic distinction of speech levels made by Tongans, this special way of using language is also characterised in terms of kinds of words: beautiful words and metaphorical words. But unlike the speech level distinctions which associate the use of words primarily with particular role relationships, there is much more here. Public performed speech is to a great extent associated with particular genres of speech, most pervasively with speeches and songs for a wide range of occasions, but also with old and new elaborately composed, practiced and performed group



song-dances for special celebrations such as the King's birthday, the opening of a new church, or a visit from an Apostle of the Church of Latter Day Saints, particularly the *ma'ulu'ulu* and the *lakalaka*. The speeches and debates of *matāpule* in formal kava circles on such occasions as funerals, weddings, and the investitures of Nobles are other important examples of genre-specific uses of rhetorical Tongan. In the Methodist denominations the literary language is also important in sermons, prayers and hymns. While the particular content and structure of these genres is specific not only to the events in which they take place, but also to the particular social groups and individuals involved, certain stylistic conventions are associated with this tradition: beautiful words, old words, rare words, metaphorical language use, proverbs, and more generally allusiveness and indirectness. In all of this, natural imagery figures prominently. Trees, rocks, and flowers index and stand for important mythical and historical events, social groups and individuals, and stories about them and relationships among them.

Traditionally the use of such language was associated with the chiefly classes. Queen Sālote is regarded as one of Tonga's greatest poets, and her songs are still sung and recorded, performed and heard on the radio. But today there are also commoners famous for their rhetorical skills. Not everyone understands all of what is said in literary language, and its instantiations are by their very nature subject to multiple interpretations, some more authoritative than others. Today then, as in earlier times, the ability to be eloquent, inspiring and entertaining in performance, and to interpret the performances of others is highly valued and conveys status upon the skilled. To perform for people honours them, and at the same time conveys honour upon the performer. In all of this, rhetorical Tongan is associated with social hierarchy, as are the speech level distinctions.

But it is Tongan speech levels that are most simply, sharply and clearly analysed by both Tongans and Western students of Tongan language and culture. I have already noted that these speech levels are conceptualised in terms of analogous words used to and about Commoners, Nobles and the King. Thus people will say that when using Tongan words for 'go', '*alu* is used to and about Commoners, *me'a* is used to and about Nobles, and *hā'ele* is used to and about the King. Or that for 'come' *ha'u* is used for Commoners, *me'a* for Nobles and *hā'ele* for the King. Several points can be made about such distinctions. First, the examples people offer to illustrate speech levels are consistently tripartite. They come in threes. Examples for which the words are the same for Commoners and Nobles, but different for the King, or the same for Nobles and the King, but different for Commoners, tend not to be given.

Second, such examples tend to background semantic relationships across these tripartite examples. For example, as is partially reflected in my two instances above, the relatively semantically unspecified commoner noun *me'a* 'thing', is used as a verb in referring to a variety of noble actions, including 'come', 'go', 'know' (*mea'i*), 'read' (*me'atohi*), 'run', 'see', 'sit' and 'speak'. But because the tripartite word lists and examples never indicate when a noble or kingly word is also a Tongan commoner word with a different meaning when used as a commoner word (rather than, say, a word borrowed from Samoan), such relationships are obscured to the outsider.

Third, these normative statements are most definitely from the point of view of the Commoner. Secondary statements, that is, those made after such initial statements, or when people are asked for more information, add that people of the same status level will use commoner terms with one another, and that the Nobles and the Royal Family use commoner words when speaking to those below them.

With regard to the lists that have been published and partially presented in Table 1, it should also be noted that each list is presented as finite, rather than open-ended, while there is in fact great

variation from list to list, and many terms identified in the Churchward dictionary (1959) as honorific or regal, or by Tongan consultants working with transcriptions of naturally occurring speech, are not included. Nor can one know from such lists that there is in fact considerable variation from speaker to speaker of the Tongan language in words that they know, use, and identify in texts as noble and kingly.

Most importantly, however, for what follows, it must be pointed out that the lists presented in Table 1 iconically convey that each level is quite distinct and separate, and analogously so. In other words, the noble level is as distinct from the commoner level as the kingly level is from the noble level.

The salient role relationships associated with the tripartite conceptualisation have changed over time, and are not unambiguous today. Thus Gifford's (1929) three categories are *muomua* 'leading chiefs', *lotoloto* 'middle chiefs' and *kakai* 'people'. Shumway (1971) titles his discussion of speech levels 'Respect Terminology for Nobility and Royalty', but then uses the Tongan terms *tu'i*, *hou'eiki* and *kakai*, which he does not gloss, but which are glossed in *Ko e Kalama* (n.d.) as 'king', 'nobles' and 'commoners'.

In everyday usage *kakai* means 'people'. *Hou'eiki* means 'chiefs'. It can also be used to refer to the present-day Nobles, but another term often used is the word borrowed from English, *nōpele*. *Tu'i* means 'king, queen or sovereign', but is used as a descriptor with other role titles to convey highest office for both traditional and modern positions, including the Tu'i Tonga. Thus common usage gives us English glosses of people, chiefs, and sovereign, while the *Ko e kalama* glosses give us the British social order hierarchy terms of Commoners, Nobles and King. There is, then, an inherent ambiguity in the relationship between the Tongan terms and their English glosses in *Ko e kalama*, a deliberate ambiguity, I think, made sense of by contemporary Tongan government-promulgated views of the relationship between Tonga as a modern nation state and the pre-nation state social order.

The present-day account given to me of the use of speech levels in pre-western contact times is that the terms associated today with the King were then used exclusively to and for the Tu'i Tonga, while the terms associated today with the Nobles were used in talking to or about chiefs. The Tu'i Tonga at the height of the Tu'i Tonga Empire in the twelfth century was regarded as embodying both sacred and secular power. Over time, as his secular power was delegated or relegated to first the Tu'i Ha'atakalaua and later to the Tu'i Kanokupolu chiefly lines, the position and the person of the Tu'i Tonga were conceptualised increasingly as of a sacred nature.

By the early 1800s the Tu'i Tonga line had considerably weakened in influence. Tāufa'āhau, who later became the first King of Tonga, Tāufa'āhau Tupou I, is credited with having unified the country through conquest and with the support of missionary influence on the British, taking the title of Tu'i Kanokupolu in the process. During the period from the 1850s to the 1870s the traditional Tongan social order was transformed into that of a nation-state, in this case a hereditary monarchy with all the symbolic and politico-legal trappings of Western European nations, particularly Great Britain, including a Constitution, and a Legislative Assembly with members from Commoner and Noble classes. This is a process well documented by Lātūkefu (1974, 1975). The Tongan Constitution of 1875, then, makes a critical legal distinction between the King, Nobles and Commoners, with regard to both rights to representation and rights to land.

At the time these events were taking place, Lātūkefu makes it clear that the Tongans aimed to establish a form of government that would be familiar to the British and other Europeans so that they

would be legitimised in British eyes and in this way stave off colonisation. They also aimed to establish a form of government that could be seen as a continuation of the pre-national Tongan social order. And they were encouraged in all of this by the Methodist missionaries upon whom Tāufa'āhau relied. Tāufa'āhau himself, as both Tu'i Kanokupolu and first King, embodied this link between the old order and the new order. And his descendents and inheritors to the throne reinforced this link by marrying spouses from the other two titled lines of the Tu'i Tonga and the Tu'i Ha'atakalaua.

I have been told that there is evidence from the analysis of written documents that in the early phases of the monarchy, which clearly was defined as a secular form of authority, and did not embody the sacred associations with the Tu'i Tonga, the terms which had been reserved for reference and address to the Tu'i Tonga were not applied to the new King. Only over time have those terms been increasingly applied to him, raising the possibility that over time the kingship has gradually taken on the aura of both sacred and secular power.

The situation regarding the transition from 'chiefly' (*hou'eiki*) to noble legitimised authority is more complex. Although Lātūkefu (1975) notes that from among the various highly prestigious chiefly lines, it was only those men who were of high status *and* had large followings who were made Nobles by the King, there are stories in Tonga today that give accounts of why certain very high and powerful chiefs of that time who should have been made Nobles were not. These same titled persons also still figure as prominently as the Nobles in some traditional activities where rank is salient.

This brief and oversimplified account of the transition to national polity helps explain the ambiguous pairing of Tongan and English words for speech levels particularly evident in *Ko e Kalama*. It should be apparent that the present tripartite analysis of speech is itself an iconic representation of the present-day political order of Tonga as a nation, yet at the same time it links this present-day political order to the past political order, giving it double legitimacy to both the outside world and the people of Tonga as they learn this representation in school.

This account will also help explain the pattern of actual usage of Tongan speech levels to which we now turn.

### 3. PRACTICE

In actual practice, the use of words identifiable and identified by Tongans as *kakai*, *hou'eiki* and *tu'i* in transcripts of Tongan discourse suggests a far less clear-cut and more complex Tongan social construction of social hierarchy within Tonga than the metapragmatic analysis embodied in the word lists.

First, and most importantly, noble and kingly terms are not just used to and about Nobles and the King.

'Kingly' terms are regularly extended in reference and address to God and Jesus, particularly in the three oldest and most traditional Methodist denominations of the Free Wesleyan Church, the Free Church of Tonga, and the Church of Tonga, but not in the more recently introduced churches such as the Church of the Latter Day Saints. Examples (1) and (2) illustrate the canonical usages of kingly speech level terms in reference to the sovereign of Tonga, and are taken from the December 1984 twentieth anniversary publication of the national newspaper, *Ko e Kalonikali Tonga* (The Tongan Chronicle).

- (1) *Ko hono ua, ko e hala 'a Kuini Sālote 'i Tisema 16 i hono ta'u 65.*  
The second was the **death** of Queen Sālote on 16 December in her 65th year.
- (2) *Pea na'e folofola 'a e Tama Tu'i 'o pehē: "Ko e Hala ki he Tu'umalie ko e ngaue malohi".*  
And the King **pointed out**, 'Hard work is the best way to riches'.

Examples (3) and (4), taken from a prayer in a Free Wesleyan Church service recorded in May 1985, illustrate the extension of kingly words to God.

- (3) *'Oku mau teiapa'a 'i ho takafalu 'i ho ao mā'oni'oni ho'o 'afio 'i he pongipongi ko eni.*  
We are huddling behind you (at your **back**) in your holy **presence** your **majesty** this morning.
- (4) *'Ei tokoni mai 'a e 'afio na, talia e lotu 'oku mau fai 'ia Sīsū Kalaisi ko ho 'alo homau Fakamo'ui.*  
Oh (yes) help us **Lord**, receive our prayers in Jesus Christ your **son** our Saviour.

In this way the pre-contact associations of these words with the sacredness of the Tu'i Tonga and the present-day authority of the King have been transferred to the Christian God. While the initial characterisation of use of these words by Tongans is always associated with the secular hierarchy, these religious usages are often mentioned second when people are asked about other uses of the speech levels. There is, however, occasional reticence about mentioning such uses, and the non-use of such extensions by some religious denominations is deliberate. Churchward's (1959) dictionary identifications of meanings always keep regal usages separate from religious usages, that is, words are never identified as entailing both, even the word for God and King, 'afio. Clearly there is ambivalence about the appropriateness of merging of secular and sacred authority semantically through such usage. That it exists and that it is widespread is, however, undeniable. At the same time, there are certain terms that are never used for both sovereign and deity, if only by virtue of logic and discourse constraints, but perhaps for other reasons for well. For example, in the same prayer cited above, the noble word for death, *pekia*, and not the kingly word *hala* used for Queen Sālote was used to refer to the death of Jesus, and there would be no other religious deaths at the top of the sacred hierarchy to refer to. There are no other major and regular extensions of kingly words that I know of outside the Royal Family.

The use of 'noble' terms is considerably more complex. Examples (5) and (6) illustrate canonical uses of noble words to refer to a noble. Example (5) is taken from the same prayer as above.

- (5) *Pea kei laumālie ai 'a Tu'ivakanō...*  
And Tu'ivakanō is still in **good health**...

Example (6) is from a *fono* (village meeting) recorded in February 1988 where a Town Officer is referring to the fact that the local Noble is leaving to go overseas to school. This example also illustrates the range of meanings of *me'a*, the commonness of its use, and the impossibility of retrieving its meanings without recourse to previous discourse context and/or shared knowledge between speaker and addressee.

- (6) *Ko e me'a 'a e 'Eiki 'o hangē ko e me'a 'a K. 'E me'a ia he ta'u e fā.*  
The **leaving** of our Noble **mentioned** by K. He **will be overseas** for four years.

Noble words are also used for non-noble chiefs. They are also extended to positions of authority in European-derived institutional complexes in Tonga. Example (7) is taken from a Magistrate's Court proceeding recorded in October 1987. The Police Prosecutor is addressing the Magistrate.

- (7) *Ko e hā ha me'a ma'olunga 'a e feitu'una 'e 'aonga ki he faka'iloa 'e fiemālie ki ai e talatalaaki.*

Any high **advice** of **yours** will be useful to the defendant and the prosecutor will be comfortable with it.

Noble words are also extended to persons in positions of authority or high status outside Tonga. Example (8) is from the same issue of *Ko e Kalonikali* as the first two examples, underscoring the way in which the Tongan sovereign has higher status than any outside authority, no matter how grand.

- (8) *Ko e 'uluaki ko e pekia 'a e taki fuoloa 'o Pilitania ko Sir Winston Churchill 'i Sānuali 24.*  
The first was the **death** on 24 January of statesman-historian Sir Winston Churchill.

Just to make matters more confusing, noble words are also used to raise the level of formality and politeness in public discourse generally. Example (9) is taken from the same Free Wesleyan church service as the examples from the prayer, but here the genre is 'announcements', for which the Tongan word *tu'utu'uni* 'instructions' is used, being spoken by the *sētuata* 'steward', rather than the preacher.

- (9) *...pea kou kole atu ko kimoutolu 'oku fai'aho mo kimoutolu 'oku me'a mai ki mu'a, 'osi mahinoia.*

...and I ask you who have birthdays and who have reasons to **come** forward, as you already know.

Here, although the noble word for 'come' is used, the noble word for 'know', which is also *me'a* is not; instead the commoner word *mahino* is used, even though this speaker uses *me'a* for 'know' in other utterances in this text. This thins out the frequency of noble words compared with the example of references to the Noble and K, his representative in example (6) from a *fono*, and suggests the way in which subtle distinctions in level of speech may be made by varying the frequency of noble words.

In example (7) I showed how a Prosecutor addressed a Magistrate. Examples (10) and (11) from the same proceeding show how the Magistrate reciprocates noble terminology to the prosecutor in (10), but does not use it to the young missionary witnesses in (11).

- (10) *'Ai mai ha'o me'a ki he hia lolotonga.*

Just **speak** (literally 'your/singular-any word') to the present crime.

- (11) *Faifekau, 'oku 'i ai ha'amo lau ki he faka'iloa?*

Missionaries, do you have something to **say** (literally 'your/dual-any word') about the defendant?

This of course shows more respect to the Prosecutor than to the witnesses.

When I asked Tongans why noble terms might be used in some of these examples, I was told it was because "a Noble might be present". This is a conventionalised response, analogous to the Magistrates' scolding of criminal defendants for the crime of using bad words (*lea kovi*) because sisters and brothers, who must not use or hear such words when co-present, might be present.

Thus, just as the levels themselves have been reanalysed and renamed in English in metapragmatic analysis to represent the highest level political order, their *use* also indicates a reanalysis of Tongan society to incorporate European-derived institutional complexes into a Tongan cultural construction of social hierarchy.



Not only do patterns of use reflect such a reanalysis, but they also reveal how much subtlety and variability there is to the ideology of social hierarchy in practice compared with the simplicity of the metapragmatic analysis represented in the word lists.

At the same time we see that the sacred associations of the words used for the Tu'i Tonga are present today in the restriction of extensions of kingly words to God and Jesus only. And we also see that in practice the traditional subtle gradings among chiefs lost in the legal creation of the category of Noble exist today in practice.

The fluidity in the use of noble terms, compared with the restrictedness of the use of kingly terms cannot be overemphasised. It is this difference that is most at odds with the metapragmatic representation of the three speech levels as analogously separate and distinct from one another. There are a number of ways of interpreting this difference, none of which I am confident in. First, it is clear that the Royal Family enjoys the respect and love of the Tongan people in a way that truly sets them apart from everyone else, which cannot be said of the Nobility, whose special status has been, since the creation of that special category, problematic. A pattern of usage that does not set them apart very clearly from the Commoners is consistent with their problematic status. But in addition, although at present I lack clear evidence of historical change in the use of these terms, a case can also be made for a general levelling of Tongan society, and for a downward movement of noble terms analogous to the downward movement of high speech levels documented by Errington (1985). Many of the noble terms listed in Shumway and *Ko e Kalama* were reported to me and observed by me to appear not just in the formal speech of Tongans, but in everyday family household use as well.

A final issue to be addressed is the extent to which the separate Tongan metapragmatic conceptualisations of the relation between social hierarchy and language use are separate in practice. Here I argue that in practice, speech levels and the distinction between everyday language use and public performed verbal art, that are kept separate in Tongan talk about language use, are not always separate in actual language use itself. In practice, performed verbal art and the noble and kingly speech levels draw upon and give effectiveness and impact to one another. They mutually influence and empower one another.

Genres of verbal art draw upon words that are perceived to be noble words to enhance the beauty and persuasiveness of what is said. In the following example (12), I quote again from the Free Wesleyan prayer used earlier, now emphasising it as a good contemporary example of rhetorical performed Tongan. The words in bold print are words identified by several Tongan informants as noble words, but as will be evident, they are not the common words in previous examples, nor are they recognised in written sources as noble. Their rarity is seen as a testimony to the knowledge and skill of the speaker.

- (12) *mau kei **kai** 'utungaki he ngaahi lelei ko ia;*  
*mau kei **tafesino**'ivai pea 'ūkuma ai 'a e ngaahi lelei katoa*  
*'oku mau pōlepole ai.*  
 we still live in **plenty**;  
 we still live in **complete** ease enjoying pleasure yours by right with no hindrance  
 of which we are proud.

I have formatted this excerpt to highlight the semantic and syntactic parallelism in these phrases to reveal their poetry-like qualities.



Reciprocally, many noble and kingly words themselves partake of the metaphorical qualities characteristic of the Tongan poetic tradition of performed verbal art. There are two general senses in which this is true. First, when we consider words that have both commoner and noble or kingly meaning, the relationship between those meanings reveals the often allusive, indirect, metaphorical and metonymic dimensions of the noble and kingly terms, dimensions we associate with poetry in many languages, including Tongan. For example, some of the kingly words associated with negative or unfortunate actions are commoner words which soften, deintensify or obscure what is being referred to. The kingly word for 'beat', which in commoner terms is *haha*, is *palai*, the commoner word for 'smudge'. The kingly word for 'drunk', which in commoner terms is *konā*, is *malahia*, the commoner word for 'misfortune from wrongdoing'. The kingly word for 'kill', which in commoner terms is *tamate'i*, is *feia*, a polite form of 'do'. The kingly word for 'spit on', which in commoner terms is *puhinga*, is *fakamokomoko*, which means 'to cool off'.

But in addition, there are some noble and kingly terms that are poetic in a more distinctly Tongan way, in that they draw upon natural imagery. For example, the commoner word for 'sky' is *langi*. *Langi* is the kingly word for 'bury', 'eye', 'face', 'head', and an element in the compound for 'eyeglasses', *langi-sio'ata*. *Langi* also has many mythical associations. The sky is the dwelling place of the gods. It is also the place from which early contact Europeans were thought to have come, as in *pālangi* 'European'. The commoner word *ma'afu* refers to a nebula configuration; it is the kingly word for 'burn'. The commoner word for 'gardenia' is *siale*, which is the root for the kingly word for 'white' *sialea*, as in white hair. In other words, these kingly words are commoner words for natural phenomena that stand for human attributes and actions when used to refer to the King, just as such natural phenomena stand for humans and their more historically noteworthy actions in the speeches of *mātapule* and the compositions of *punake*.

In sum, the different expressions of the cultural constitution of social hierarchy in Tonga are kept analytically distinct, and quite often they are kept separate in practice. It cannot be said, for example, that the newspaper writer, the Magistrate, or the Town Officer have been poetic in my examples of their use of noble terms.

However, these special forms of language use also merge, interpenetrate, and draw on one another for their power and effectiveness in many aspects of their expression.

#### 4. CONCLUSION

For Tongans, talk about the existence and nature of Tongan speech levels and actual use of vocabulary identified by Tongans as associated with three distinct social categories are closely related but distinguishable ways of culturally constructing social hierarchy in Tonga.

Both have undergone change as the dominant interpretive framework for social hierarchy in Tonga has made a transition to nation-state. In the talk about practice, or metapragmatics of speech levels, there has been a transformation of the key social identities associated with the three levels from people, chiefs and the Tu'i Tonga to Commoners, Nobles and Kings. But at this level of analysis, only the key constitutional categories of Tongan modernity are in evidence. The continuing sacred character of Tongan kingly words manifest in practice is, however, obscured in the metapragmatic word lists. And the multiple and fluid sources of traditional and modern authority indexed by noble words manifest in practice cannot be retrieved from the rigidly tripartite analysis. Nor is it possible

to recognise from the word lists the way in which speech levels and the performance tradition are intertwined and draw upon one another, even as they are truly distinct in some ways.

Tongan metapragmatic analysis of speech levels is itself very Tongan as I see it. It is one of many ways the Tongans have of presenting the same model of their entire society to outsiders and to themselves: overtly and expressly hierarchical, comprehensible to Europeans, using European terms (king, noble, commoner), yet clearly linked to the Tongan pre-contact past in such a way that each legitimises the other.

At the same time, their metapragmatic analysis has much in common with the ways in which people who speak other languages in other cultures use and interpret their own languages and cultures. First, it focuses on the segmentable phenomena of words, which are more accessible to analysis than non-segmental aspects of language structure. Second, some particular aspects of practice are carved out and highlighted from all that is experientially merged. One could never know, for example, of the very strong egalitarian values in Tongan culture from this analysis, except indirectly from some of the extensions of noble terminology. Third, such constructs normativise or idealise linguistic phenomena which are more variable and complex behaviourally than the constructs suggest, as we have seen in language variationist research in the Labovian tradition. For example, Charles Ferguson argued some years ago that many languages have high and low varieties, a situation which he referred to as diglossic. These varieties were said to differ in both form and function in similar ways in unrelated languages. Yet recently diglossic varieties have been shown to be not sharply distinguished in linguistic form, but rather as differing in terms of the frequency of particular forms, that is, as variably rather than categorically different. And finally, such constructs idealise not only linguistic phenomena, but also sociocultural phenomena, not just overgeneralising, but also projecting a (self) image of society that gives salience to certain aspects of social ordering over others.

## BIBLIOGRAPHY

- BLIXEN, Olaf, 1966, Lenguaje honorifico y comportamiento reverente en Samoa y Tonga. *Comunicaciones Antropológicas del Museo de Historia Natural de Montevideo* 6:1-39.
- BAUMAN, Richard, 1975, Verbal art as performance. *American Anthropologist* 77/2:290-311.
- BOAS, Franz, 1911, Introduction. In F. Boas, ed. *Handbook of American Indian languages*, Pt. 1, 1-83. Bureau of American Ethnology, 40. Washington, D.C.
- CHURCHWARD, C. Maxwell, 1953, Tongan grammar. London: Oxford University Press.
- 1959, Tongan dictionary. London: OUP. Tonga: Government Printing Press.
- COLLOCOTT, E.E.V., 1925, Supplementary Tongan vocabulary. *Journal of the Polynesian Society* 34/2:146-169; 193-213.
- DURANTI, Alessandro, In press, Language in context and language as context: the Samoan respect vocabulary. To appear in C. Goodwin and A. Duranti, eds *Rethinking context*. Cambridge: Cambridge University Press.
- ERRINGTON, J. Joseph, 1985, On the nature of the sociolinguistic sign: describing the Javanese speech levels. In E. Mertz and R.J. Parmentier, eds *Semiotic mediation*, 287-310. New York: Academic Press.
- GIFFORD, E.W., 1929, *Tongan society*. B.P. Bishop Museum Bulletin 8. Honolulu.
- KAEPPLER, Adrienne, 1976, Dance and the interpretation of Pacific traditional literature. In A.L. Kaeppler and H.A. Nimmo, eds *Directions in Pacific traditional literature*, 195-216. Honolulu: Bishop Museum Press.
- Ko e Kalonikali Tonga, 1984, Twentieth anniversary issue, December.
- Ko e Kalama 'i he Lea Faka-Tonga, n.d., Printed by the Free Wesleyan Church Education Office.

- LĀTŪKEFU, Sione**, 1974, *Church and state in Tonga: The Wesleyan Methodist missionaries and political development*. Canberra: Australian National University Press.
- 1975, *The Tongan Constitution: a brief history to celebrate its centenary*. Nuku'alofa: Tonga Traditions Committee Publications.
- MARCUS, George**, 1980, *The nobility and the chiefly tradition in modern Tonga*. Honolulu: University of Hawaii Press.
- MILNER, G.B.**, 1961, The Samoan vocabulary of respect. *Journal of the Royal Anthropological Institute* 91/2:296-317.
- SHIBAMOTO, Janet S.**, 1987, The womanly woman: manipulation of stereotypical and nonstereotypical features of Japanese female speech. In S. Philips, S. Steele and C. Tanz, eds *Language, gender and sex in comparative perspective*, 26-49. Cambridge: Cambridge University Press.
- SHUMWAY, Eric B.**, 1971, *Intensive course in Tongan*. Honolulu: University of Hawaii Press.
- 1977, Ko e fakalāngilāngi: the eulogistic function of the Tongan poet. *Pacific Studies* 1/1:25-34.
- SILVERSTEIN, Michael**, 1979, Language structure and linguistic ideology. In P.R. Clyne, W.F. Hanks and C.L. Hofbauer, eds *The elements: a parasession on linguistic units and levels*, 193-247. Chicago: Chicago Linguistic Society.
- 1981, The limits of awareness. *Working Papers in Sociolinguistics*, Numbers 81-87. Austin: Southwest Educational Development Laboratory.
- TALIAI, 'Opeti Manisela**, 1989, Social differentiation of language levels in Tonga. MA thesis, University of Auckland.
- TREGEAR, Edward**, 1895, Ceremonial languages. *New Zealand Institute Transactions, Royal Society of New Zealand* 27:593-597.
- WATSON-GECEO, Karen and David W. GECEO**, 1986, Calling-out and repeating routines in Kwara'ae children's language socialization. In B. Schieffelin and E. Ochs, eds *Language socialization across cultures*, 19-50. Cambridge: Cambridge University Press.

FINAL VOWEL LENITION IN MICRONESIAN LANGUAGES:  
AN EXPLORATION OF THE DYNAMICS OF DRIFT

KENNETH L. REHG

1. OVERVIEW

All Micronesian languages, to varying degrees, exhibit lenition of word-final or phrase-final vowels. By lenition, I mean the weakening of final vowels by processes which result in a decrease in their sonority. The extreme case of weakening is, of course, deletion. Micronesianists commonly refer to this phenomenon of final vowel lenition as *erosion from the right* – a term first employed by Bender (1969:43). With reference to final vowel apocope, he observes:

Quite common in Micronesian languages is...‘erosion from the right’ which causes stem final vowels – the final vowels of CVCV stems – to be dropped unless they are ‘protected’ by suffixes.

Jackson (1983:432), too, notes that, among thirteen phonological innovations shared by all Micronesian languages, one is what he more generally characterises as “the loss of final vowel information”.

This paper explores the loss of final vowel information in Micronesian languages and presents arguments for three claims about this phenomenon. (1) While all Micronesian languages show final vowel lenition, the processes resulting in such lenition are not identical among all the languages. (2) Because different processes result in parallel developments, final vowel lenition constitutes an example of drift. (3) Drift, in this instance, can plausibly be explained as a consequence of the retention, extension, or suppression of a constellation of phonological processes present in Proto Micronesian.

2. HISTORICAL RELATIONSHIPS OF MICRONESIAN LANGUAGES

The term ‘Micronesian language’ is employed here to refer to any of the various languages that Bender (1971) classifies as ‘Nuclear Micronesian’. These include Kiribati (formerly Gilbertese), Marshallese, Kosraean (formerly Kusaian) and the languages of the Trukic and Ponapeic continua. The Trukic continuum includes, among other languages, Lagoon Trukese, Mortlockese, Satawalese, Woleaian and Ulithian. Ponapeic includes Ponapean, Mokilese and Pingelapese. All of these

languages, possibly together with Nauruan, have previously been argued to form a Micronesian (MIC) subgroup within Oceanic (Bender 1971; also Jackson 1983).<sup>1</sup>

Trukic and Ponapeic are well-defined subgroups within MIC.<sup>2</sup> Whether or not further subgrouping is possible within MIC is a matter of some controversy. A view of the relationships among these languages held by some Micronesianists working at the University of Hawaii in the late 1970s is represented by the following tree.<sup>3</sup>

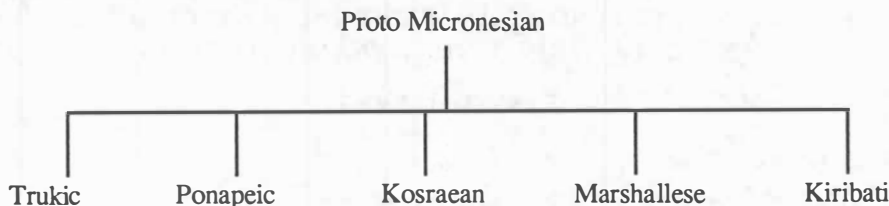


FIGURE 1: HISTORICAL RELATIONSHIPS WITHIN MICRONESIAN (FIVE COORDINATE BRANCHES)

In this figure, Trukic, Ponapeic, Kosraean, Marshallese and Kiribati represent five coordinate branches of Proto Micronesian (PMC). Implicit in this view is an assumption that the settlement of those islands or island groups not part of the geographic homeland of PMC must have taken place with sufficient rapidity that further subgrouping could not have developed.

Jackson (1983), however, argues against this representation of historical relationships. The following figure (somewhat simplified) represents his claims.<sup>4</sup>

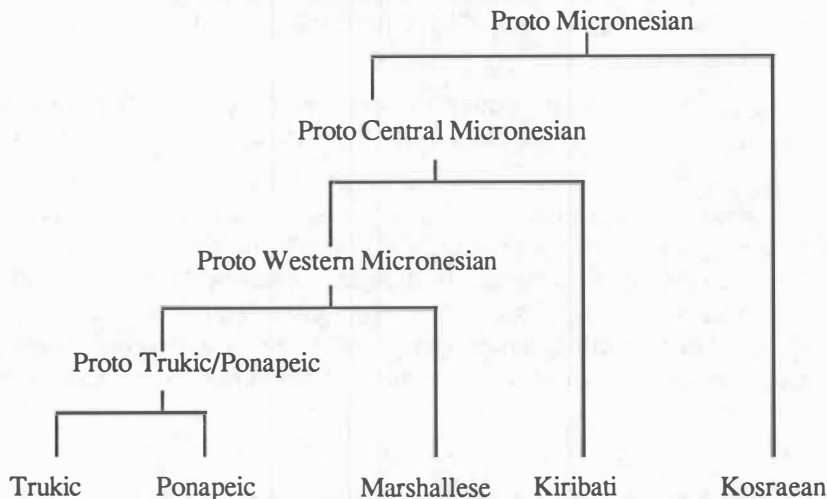


FIGURE 2: HISTORICAL RELATIONSHIPS WITHIN MICRONESIAN (POSSIBLE SUBGROUPS)

<sup>1</sup>Nauruan is not dealt with in this paper due to a paucity of reliable data. This language, though, also exhibits lenition of final vowels.

<sup>2</sup>See Jackson (1983) and Rehg (1981).

<sup>3</sup>I refer here to the regular participants in what came to be known as the Micronesian Seminar; these are the individuals listed in Bender et al. (1983). In retrospect, it is difficult to recall precisely who held this view at what point in time. Prior to Jackson (1983), however, such a tree represented a relatively uncontroversial claim about relationships within MIC.

<sup>4</sup>Jackson's tree (1983:433) further defines relationships within Trukic and suggests a closer relationship may exist between Ponapean and what he terms Central Trukic.

Jackson (1983:437-440) outlines several possible settlement scenarios that would account for this tree, all of which preclude the more rapid pattern implicit in Figure 1.

Commentary on how the subject of this paper bears on this issue is provided in section 9.

3. THE SCOPE OF FINAL VOWEL LENITION WITHIN MICRONESIAN

Before attempting to provide an account of the historical development of final vowel lenition, the scope of this phenomenon within MIC must first be considered. Data are drawn from six representative languages – Kiribati (KIR), Marshallese (MRS), Kosraean (KSR), Ponapean (PNP) from the Ponapeic group, and from the Trukic group, Lagoon Trukese (TRK) and Woleaian (WOL). These languages were selected for consideration because they exemplify the range of final vowel lenition processes that occur in MIC languages.<sup>5</sup>

3.1 FULL APOCOPE

Final vowel lenition in most MIC languages manifests itself as what I will term ‘full apocope’. Full apocope, in contrast to ‘partial apocope’ (discussed below), results in (1) the loss of word-final short vowels and (2) the shortening of word-final long vowels. The following data, contrasting PMC reconstructions with synchronic forms from Ponapean, are illustrative of full apocope.<sup>6</sup> In these examples, long vowels are written doubled. The vowel length that occurs in the Ponapean forms in the first column is discussed in Rehg (1981, 1984b).

DELETION OF FINAL SHORT VOWEL			SHORTENING OF FINAL LONG VOWEL		
PMC	<i>*kuli</i>	skin	PMC	<i>*pii</i>	vagina
PNP	<i>kiil</i>	skin	PNP	<i>(pii)pi</i>	vagina
PMC	<i>*umwu</i>	earth oven	PMC	<i>*fituu</i>	star
PNP	<i>uumw</i>	earth oven	PNP	<i>usu</i>	star
PMC	<i>*rongo</i>	hear	PMC	<i>*wuaa</i>	fruit
PNP	<i>roong</i>	news	PNP	<i>wwa</i>	fruit

In Ponapean, phonologically predictable reflexes of PMC final vowels surface when suffixes follow. The following examples illustrate the forms above in combination with the construct suffix *-n* ‘of’.<sup>7</sup>

PNP	<i>kilin</i>	skin of	PNP	<i>piipiin</i>	vagina of
PNP	<i>umwin</i>	earth oven of	PNP	<i>usuun</i>	star of
PNP	<i>rongen</i>	news of	PNP	<i>wwaan</i>	fruit of

Given the synchronic alternations that occur in Ponapean, one can, within the framework of generative phonology, posit the following underlying representations for these forms.

<sup>5</sup>Unless otherwise specified, sources of data for these languages are: Kiribati (Harrison n.d.); Marshallese (Bender 1968); Kosraean (Lee 1975); Ponapean (Rehg 1981); Lagoon Trukese (Sugita n.d.) and Woleaian (Sohn 1975).

<sup>6</sup>PMC reconstructions are from Bender et al. (1983).

<sup>7</sup>For further discussion, see Rehg (1984a, 1986).



## SYNCHRONIC BASE FORMS

<i>kili</i>	skin	<i>piipii</i>	vagina
<i>umwi</i>	earth oven	<i>usuu</i>	star
<i>ronge</i>	hear	<i>wwaa</i>	fruit

The deletion of final short vowels can be accounted for by the following rule.

## FINAL SHORT VOWEL APOCOPE

$$V \text{ ---} \rightarrow \emptyset / VC_1 \_\_\_\#$$

This rule deletes the final short vowel of a minimally disyllabic stem. If long vowels are represented as underlying geminates, as in the examples above, then the shortening of final long vowels can be accounted for by the next rule.

FINAL LONG VOWEL APOCOPE<sup>8</sup>

$$V \text{ ---} \rightarrow \emptyset / V \_\_\_\#$$

These two rules can obviously be conflated, as follows.

## FINAL VOWEL APOCOPE

$$V \text{ ---} \rightarrow \emptyset / VC_o \_\_\_\#$$

Other languages that have apparently undergone full apocope are Lagoon Trukese, Marshallese and Kosraean. In the case of Lagoon Trukese, a synchronic rule of final vowel apocope identical to that for Ponapean has been formulated, first by Irwin Jay Howard (pers. comm.), and subsequently by other students of this language (Bender 1973). Precisely how apocope has affected Marshallese and Kosraean is less clear.

In Marshallese, all PMC final short vowels undergo apocope and all final long vowels surface as short vowels followed by a glide, as illustrated by the following examples.

PMC	<i>*kuli</i>	skin	PMC	<i>*fituu</i>	star
MRS	<i>kil</i>	skin	MRS	<i>yijiw</i>	star

As in Ponapean, morphophonemic alternations justify treating short vowel apocope as a synchronically productive rule. It is less certain whether or not Marshallese ever had a rule of long vowel apocope. Synchronically, all long vowel sequences in this language are analysed by Bender (1968) as  $V_1GV_1$  sequences, where G represents a glide. The underlying representation for 'star' is thus /yijiwi/. Since within this analysis there are no long vowels, the rule of short vowel apocope suffices to account for surface synchronic forms. Further, if the rephonemicisation of PMC long vowels took place early in the history of Marshallese, before final vowel deletion, then it is possible that a rule of long vowel apocope never existed. But, by whatever route, full apocope is exhibited in this language.

In Kosraean, morphophonemic alternations apparently do not justify establishing synchronic underlying representations that exhibit reflexes of PMC final short vowels (Lee and Wang 1984; see also section 5). It is clear, however, that a rule of short vowel apocope operated diachronically in this language, as illustrated by the following examples.

<sup>8</sup>This rule also correctly predicts that the final vowel of a non-identical sequence will delete; cf. POC *\*mpua* 'areca nut', PNP *pwu*: 'areca nut'.

PMC	* <i>kutu</i>	louse	PMC	* <i>wulunga</i>	pillow
KSR	<i>kuut</i>	louse	KSR	<i>ilung</i>	rest, prop

Whether or not final long vowels were shortened is less obvious. In Kosraean, only non-low vowels occurring in the first syllable of polysyllabic roots may occur short (Lee 1975:32).<sup>9</sup> All other vowels are long. The issue of whether a rule of long vowel apocope ever existed is thus obscured by this phonological innovation, unique among MIC languages. It is consequently possible that long vowels were never shortened in Kosraean and that this language did not undergo full apocope. It should be noted, however, that all other MIC languages that have undergone short vowel apocope have also shortened final long vowels. Therefore, one might reasonably presume that this language did undergo full apocope, but that later developments resulted in the lengthening phenomena that are currently exhibited. Further comment on this issue is provided in section 7.2.

### 3.2 PARTIAL APOCOPE AND DEVOICING

Two of the six languages under consideration here – Woleaian and Kiribati – do not exhibit full apocope, but instead synchronically both delete and devoice final vowels under the conditions described below.

Woleaian is the simpler case. In this language (1) phrase-final short vowels are devoiced, and (2) phrase-final long vowels are shortened.<sup>10</sup>

Lenition in Kiribati is more complex. Final vowel apocope affects only high vowels after nasals. Two such developments are evidenced in this language. (1) Word-final short high vowels (/i/ and /u/) are deleted when preceded by a nasal. (2) Phrase-final long vowels are optionally shortened when preceded by a nasal. Harrison (n.d.), however, notes there are in fact many exceptions to rule 1. As further discussed in section 4.1, he argues that this rule, unlike rule 2, is no longer fully productive.

Kiribati also devoices final vowels under the following conditions: (1) Phrase-final short high vowels obligatory devoice after /t/ (phonetically [s]). (2) Phrase-final short high vowels optionally devoice after other consonants.<sup>11</sup> (3) Phrase-final short non-high vowels optionally devoice after geminate nasals; mid-vowels are more likely to do so than low vowels.

### 3.3 SUMMARY OF RULES

A summary of the rules of apocope and devoicing in the MIC languages under consideration is provided here. Full apocope is represented by the two rules presented in section 3.1, rather than by the single collapsed version of these rules. All rules are assumed to operate synchronically unless preceded by (d), indicating that the rule is attested only diachronically; (d?) signals that the diachronic status of the rule is unclear. A question mark in parentheses indicates uncertainty regarding whether

<sup>9</sup>Lee (1975:32) actually states: "...a short vowel can occur only in the first syllable of a word that is made up of more than two syllables". The data he cites, however, makes it clear that he meant to say 'two or more syllables'.

<sup>10</sup>Sohn (1975:18-22) describes these phenomena as occurring in word-final position. However, other of his comments indicate that he is in fact referring to phrase-final position; it is explicitly characterised as such in Tawerilmang and Sohn (1984:178). Jackson (1983) notes that Pulu Ana, Sonsorolese and Tobi pattern like Woleaian. All other members of the Trukic group undergo full apocope. Jackson also claims that final short vowels have been lost relatively recently in at least some Trukic languages. "Elbert (1974) and Sohn and Bender (1973) have observed the voiceless retention of \*o in PUL [Puluwat] and ULI [Ulithian] respectively, especially after a velar consonant" (Jackson 1983:218).

<sup>11</sup>Stephen Trussel, personal communication.

the rule ever existed in the language. Optional rules are marked by a preceding (o). A double vertical stroke represents phrase boundary and N represents [+nasal].

## APOCOPE OF SHORT AND LONG VOWELS

LANGUAGE(S)	SHORT VOWELS	LONG VOWELS
PNP, TRK	V---> Ø/VC <sub>1</sub> ___#	V---> Ø/V___#
MRS	V---> Ø/VC <sub>1</sub> ___#	(d?) V---> Ø/V___#
KSR	(d) V---> Ø/VC <sub>i</sub> ___#	(?) V---> Ø/V___#
WOL	--	V---> Ø/V___
KIR	V---> Ø/VN___# +hi	(o) V---> Ø/NV___   +hi

## DEVOICING OF SHORT VOWELS

WOL	V---> [-vd]/VC <sub>1</sub> ___	
KIR	V---> [-vd]/Vt___   +hi	
KIR	(o) V---> [-vd]/VC <sub>1</sub> ___   +hi	
KIR	(o) V---> [-vd]/VNN___   -hi	(especially mid V's)

## 4. PREVIOUS ANALYSES OF FINAL VOWEL LENITION IN PROTO MICRONESIAN

When attempting to understand the history of MIC languages, the facts of final vowel lenition noted above are a source of perplexity. Obviously, one cannot attribute full apocope to PMC, since, apart from the uncertainty of the Marshallese and Kosraean data, neither Kiribati nor Woleaian exemplify this phenomenon. Further, one cannot assume the existence of full apocope in some language ancestral to all MIC languages except Kiribati and Woleaian, since there is no evidence for such a subgroup. It is also certain that Kiribati does not exclusively subgroup with Woleaian. Woleaian is indisputably a Trukic language and Kiribati is not (see section 2). Moreover, while Kiribati and Woleaian both delete and devoice vowels, the rules governing these phenomena in the two languages are different. Finally, one may note that the islands of Kiribati and Woleai are located at the eastern and western extremities of Micronesia and are separated by approximately 3,000 miles of ocean.

The obvious question, then, is what happened in the history of these languages that resulted in this peculiar distribution of rules of final vowel lenition. Two possible answers have been suggested by Marck and Jackson.

## 4.1 MARCK

Marck (1977) speculates that Kiribati might once have had a more general final vowel devoicing rule and that the redevelopment of final voiced vowels might have been a result of Polynesian

influence from Tuvalu. Given such a scenario, it is thus reasonable to attribute a rule of final short vowel devoicing to PMC. Because the subsequent and independent loss of these devoiced vowels in most MIC languages could then be viewed as a natural and expected phonological development, the problem of the distribution of rules of final vowel lenition in the daughter languages would be solved.

This is not a totally impossible scenario, but neither is it a satisfyingly explanatory one. As Jackson (1983:322) has noted, it is by no means obvious how one might falsify such a claim. One may also question, if Kiribati did once have a rule devoicing all final vowels, just what was the nature of the proposed Polynesian influence? Did all vowels except high vowels after nasals revoice, or did all vowels revoice before the development of the rule of short vowel apocope attested synchronically? Marck's conjecture seems, in fact, to pose more problems than it was originally intended to solve. This, however, is not to discredit his more basic observation that Polynesian influence may have played some role in the history of Kiribati phonology. Indeed, Harrison (n.d.) suggests that the short vowel apocope rule of Kiribati previously described may have ceased to be fully productive at the time of Polynesian contact; therefore, final short vowels after nasals were not deleted in Polynesian borrowings, nor possibly in lexical items coined subsequent to this era.

#### 4.2 JACKSON

Jackson (1983) provides two comments relevant to the conditions present in PMC that might have led to final vowel apocope in the daughter languages. In one instance (1983:219), he notes: "...there is some evidence in Kiribati that final short high vowels may have been devoiced after sonorants as early as Proto-Micronesian". He also notes in the same discussion that "all Micronesian languages regularly shorten final long vowels before pause, suggesting that such a shortening rule probably also applied in Proto-Micronesian". In a subsequent section (1983:322) he comments: "there is a likelihood that some loss of final vowel information occurred in PMC. At the least, we can probably assume some phonetic shortening of word-final long vowels, and it seems very probable that final \*i (if not all high vowels) was devoiced as well".

Jackson's observations, while fundamentally sound, are not without problems. First, with regard to his conjecture about PMC final short vowels, one may note that Kiribati does *not* present evidence that final short high vowels might have devoiced after sonorants; Kiribati only supports the position that high vowels may have devoiced after nasals. Second, with respect to his (possibly contradictory) claims about final long vowels, one may observe that it is not precisely the case that all MIC languages shorten final long vowels before pause. Kosraean exhibits no such rule synchronically and Kiribati *optionally* shortens final long vowels before pause only if they are *high* and after *nasals*. His claim that PMC word-final long vowels might have undergone a degree of 'phonetic shortening', while probably correct, is a matter of speculation. Most importantly, however, Jackson's claims fail to take into account the important observation that, in identical environments, final short vowel lenition must have preceded (or co-occurred with) the lenition of corresponding final long vowels, as observed in both Kiribati and Woleaian. If this were not the case – if long vowel shortening preceded short vowel lenition – then surface contrasts of length would be neutralised. This possibility is illustrated by the hypothetical forms below, where vowels after undergoing lenition are represented by the letter *v*.

Base forms	...CV	...CVV
Long Vowel Shortening	--	...CV
Short Vowel Lenition	...Cv	...Cv

No MIC language, however, has neutralised length contrasts in final vowels. Final long vowel shortening, then, must have developed subsequent to (or less likely, simultaneously with) the devoicing or deletion of final short vowels. Consequently, Jackson's claims require further refinement, as further discussed in section 7.2.

## 5. PROCESSES RESULTING IN FINAL SHORT VOWEL LENITION

The developments leading to final vowel lenition in MIC cannot, in any reasonably satisfactory manner, be accounted for by the hypotheses of either Marck or Jackson. In fact, the problem of accounting for such lenition is considerably more complex than either of their scenarios would suggest. While it appears true that, under various conditions, final long vowels simply shortened, the processes leading to the lenition of final short vowels are more varied. To support this claim, it is necessary first to consider what I shall term here 'recovered' vowels in each of the MIC languages under consideration, and then to examine the types of processes that apparently led to the facts observed.

### 5.1 RECOVERED VOWELS

PMC, like Proto Oceanic (POC), had a conventional five vowel system – /a, e, i, o, u/. In PMC, these vowels occurred both short and long.<sup>12</sup> Among the daughter languages, only Kiribati synchronically reflects such a system. Other MIC languages have either reduced or expanded this inventory. In some languages that exhibit apocope of short base-final vowels, however, the inventory of what I shall term here 'recovered' vowels – those base-final short vowels that surface only before suffixes – is a subset of the total number of vowel phonemes in the language.

The following descriptions summarise the relevant facts for each of the languages discussed in this study. Unless otherwise noted, the inventory of long vowels in each language is identical to the inventory of short vowels.

#### (1) Kiribati

Short vowel inventory: /i, u, e, o, a/

Rule of apocope: V ---> Ø/VN\_\_\_#  
+hi

Recovered vowels: /i/

Comment: Word-final \*/i/ and \*/u/ merge as /i/ after nasals.

#### (2) Marshallese

Short vowel inventory: /i, ê, e, a/

Rule of apocope: V ---> Ø/VC1\_\_\_#

Recovered vowels: All

Comment: Note that all vowels have been fronted to a single palatal series in Marshallese.

<sup>12</sup>Some PMC reconstructions (Bender et al. 1983) also include *u'*, an allophone of /u/, as discussed in section 7.4.



## (3) Kosraean

Short vowel inventory: /i, ī, u, e, ə, o, ε, ʌ, ɔ, æ, a, oa/

Rule of apocope: V---> Ø/VC<sub>1</sub>\_\_\_\_#

Recovered vowels: None

Comments: (1) As previously noted in section 3.1, low vowels are always long in Kosraean. (2) Wang (Lee and Wang 1984:436) notes: "There is no synchronic evidence for retention of historical [base-final short] vowels, even in underlying forms. The variation one is apt to encounter within a single possessive paradigm, as well as the competition between alternative paradigms for the same noun, both testify to the present fluctuating state of KSR morphophonemics".

## (4) Ponapean (Northern Dialect)

Short vowel inventory: /i, u, e, o, ε, ɔ, a/

Rule of apocope: V---> Ø/VC<sub>1</sub>\_\_\_\_#

Recovered vowels: /i, ε/

Comments: (1) In Rehg (1986), I argue that base-final /a/ has been reanalysed as the initial segment of following suffixes. (2) /e/ arises from earlier /...εCi/ sequences; it never occurs in base-final position. (3) Final short back vowels were fronted in Ponapean (Rehg 1984a). /a/ is a front vowel, as discussed in Rehg (1986).

## (5) Lagoon Trukese

Short vowel inventory: /i, ī, u, e, ə, o, æ, a, ɔ/

Rule of apocope: V---> Ø/VC<sub>1</sub>\_\_\_\_#

Recovered vowels: /i, ī, u, a/

Comment: Final short mid-vowels were raised to high vowels.

## (6) Woleaian

Short vowel inventory: /i, u, u, e, o, a/

Rule of apocope: None

Recovered vowels: None, since apocope does not apply to short vowels.

Comments: (1) Sohn (1975:18) notes that Woleaian has an inventory of eight long or 'doubled vowels'. In addition to long counterparts of the six short vowels, /e:/ (a rounded mid-central vowel) and /o:/ also occur. (2) Final short /a/ is devoiced and raised to [e] or [o] by rules described by Sohn (1975:26-27).



## 5.2 SHORT VOWEL LENITION PROCESSES

The preceding observations about vowel systems in MIC support the position that short vowel apocope was almost certainly preceded by one or more of three distinct processes – devoicing, fronting and raising – all of which represent types of lenition or weakening.<sup>13</sup> Therefore:

- (1) Voiceless vowels are weaker than voiced vowels.
- (2) Front vowels are weaker than back vowels.
- (3) High vowels are weaker than low vowels.

One or more processes leading to such lenition were or are operative in all MIC languages.

Devoicing of final short vowels, as previously noted, occurs in both Kiribati and Woleaian. It is likely that devoicing occurred in other MIC languages as well, as a precursor to apocope, but it is impossible to verify this claim based upon synchronic evidence.

Fronting of final short vowels is evidenced by Kiribati, where final /u/ after nasals fronted to /i/, by Marshallese, where all vowels have been collapsed into a single front or palatal series, and in Ponapean, where all final back vowels were fronted. No synchronic evidence for fronting occurs in other MIC languages.

Raising of final short vowels is exhibited only by Trukic languages. In Lagoon Trukese, final short mid-vowels were raised to high vowels, and in Woleaian, as a result of a productive rule, short /a/ raises to either [e] or [o], as previously noted.

Because Kosraean does not exhibit reflexes of final short vowels, it is unclear which of these processes might have been operative in this language.

It is obvious, then, that final short vowels in MIC languages neither undergo nor have undergone identical processes of lenition. Further note that, since final short vowel lenition must have taken place prior to or coincident with the shortening of final long vowels (section 4.2), it is consequently not possible to attribute the shortening of *all* final long vowels in MIC languages to a single ancestral rule.

## 6. FINAL VOWEL LENITION AS AN INSTANCE OF DRIFT

Based upon the variety of observations made thus far – the historical relationships of these languages, the scope of vowel lenition processes, and the processes that play a role in such lenition – it seems reasonably clear that the loss or lenition of final vowels in MIC languages must, to some extent, have involved independent developments.

One might account for such independent developments in several ways. One possibility, of course, is that the presence of rules of final vowel lenition in all of these languages is simply a result of accident. Or, one may prefer to argue that lenition in MIC is a result of convergence, the basic rationale here being that if a design feature of a language is good enough to have evolved once, then it is good enough to evolve more than once, from different starting points, in multiple settings. Neither of these explanations, however, is entirely satisfactory. Both fail to explain why *all* MIC languages, in at least some environments, shorten final long vowels (in the case of Kosraean, diachronically) and

<sup>13</sup>See Foley (1977:44) for arguments in support of claims (2) and (3).

weaken final short vowels, whereas *no* Polynesian languages, for example, exhibit lenition of comparable scope.<sup>14</sup>

A preferable explanation, then, might be one that invokes the notion of diffusion; therefore, the loss of final vowel information in all of these languages might be accounted for by assuming that such loss originated in some innovating core language and that this phenomenon spread to all other languages. Such speculation cannot be disproved, but arguing the plausibility of such a possibility is by no means an easy undertaking. The languages under consideration here are spoken on small islands spread out over an area larger than the continental United States. Further, while there is evidence for at least some contact between these island groups, there is, currently at least, no evidence for extensive borrowings among all of these languages.<sup>15</sup> The question then arises as to how much contact, and under what circumstances, is required for one or more phonological rules to diffuse? Obviously, phonological rules are not viral infections that insidiously spread from host to host. Instead, they are behaviours that may or may not be imitated by other speakers. Why should final vowel lenition be imitated by speakers of *all* MIC languages?

Still another possible explanation, and the one that will be argued here to be preferable, is that the development of final vowel lenition in the daughter languages of PMC constitutes an example of 'drift', a phenomenon viewed in this paper as a type of natural linguistic evolution.

The term 'drift', unfortunately, carries many negative connotations, and perhaps too often is employed as a term more appropriate to metaphysics than to linguistics. Sapir (1921:183), for example, characterises drift as a result of the "psychic undercurrents" of a language. He uses the term primarily to refer to the direction of change that one can observe in a single language over time, but he also notes (1921:172): "The momentum of the more fundamental...drift is often such that languages long disconnected will pass through the same or strikingly similar phases". Lakoff (1972:178) defines drift as a "metacondition on the way the grammar of a language *as a whole* will change". It is not clear to me, however, what the empirical correlates are of either "psychic undercurrents" or of "metaconditions".<sup>16</sup>

A more substantive characterisation of drift is provided by Bynon (1977:250) who notes that "parallel development after separation should...be sought in the joint inheritance of *intrinsic structural properties...*". In this view, independent parallel developments of the type observed in MIC languages, like all evolutionary phenomena, are a result of chance, but not of accident. It is this view of drift that will subsequently be explored.

## 7. PHONOLOGICAL PROCESSES PRESENT IN PROTO MICRONESIAN

Given Bynon's conception of drift, the issue now to be explored is what intrinsic structural properties might have been present in PMC that resulted in the lenition of final vowels in all the daughter languages? The position taken in this paper is that an attempt to answer this question necessarily involves the reconstruction of at least a portion of the phonological component of PMC.

<sup>14</sup>Final and medial vowel devoicing does, of course, take place in some Polynesian languages. For example, Tongan optionally devoices unstressed vowels; Samoan devoices final short high vowels after voiceless consonants, and Kapingamarangi devoices all final short vowels.

<sup>15</sup>This is not to exclude the possibility of *any* borrowing among these languages. See, for example, Rehg and Bender (1990).

<sup>16</sup>Lakoff, herself, finds the concept of a 'metacondition' puzzling. She admits: "It is...not clear what sort of thing this metacondition is..." (Lakoff 1972:179). For a critique of Lakoff and a vindication of Sapir, see Vennemann (1975) and Blust (1978). Insightful discussions of drift are also provided in Donegan and Stampe (1983) and in Blust (1990).

That is, the crucial structural properties of PMC that gave rise to such widespread lenition apparently lie not in its phonemic inventory, but rather in the phonological rules that must have been operative upon this inventory. Rules that can be reconstructed in five areas of PMC phonology are examined here.

### 7.1 PENULTIMATE STRESS

Substantial evidence, not considered here, indicates that PMC, like POC, exhibited penultimate stress.<sup>17</sup> Final short vowels were thus unstressed and subject to further weakening. Final long vowels might then have shortened simultaneously with or subsequent to the lenition of final short vowels, since surface contrasts between short and long vowels would not then be neutralised.

### 7.2 FINAL VOWEL LENITION

If one assumes that devoicing of final short high vowels preceded their deletion, then one may argue that all the daughter languages of PMC minimally attest lenition of the scope defined by the following two rules; these rules, ordered as presented below, may thus be reconstructed for PMC.

#### FINAL SHORT VOWEL DEVOICING

$$\left[ \begin{array}{l} +\text{syl} \\ +\text{high} \\ -\text{long} \end{array} \right] \quad \text{--->} \quad [-\text{vd}]/[+\text{syl}] [+nas] \_\_\_\_ \parallel$$

Phrase-final short high vowels are devoiced after a nasal.<sup>18</sup>

#### FINAL LONG VOWEL SHORTENING (OPTIONAL)

$$\left[ \begin{array}{l} +\text{syl} \\ +\text{high} \\ +\text{long} \end{array} \right] \quad \text{--->} \quad [-\text{long}]/[+\text{nas}] \_\_\_\_ \parallel$$

Phrase-final long high vowels optionally shortened after a nasal and before pause.

Note that these rules are refinements of Jackson's claims discussed in section 4.2 and that the second rule may be reconstructed for PMC only if Kosraean did undergo full apocope, as discussed in section 3.1. If final long vowels never shortened in Kosraean, then one must assume either that the second rule originated in Jackson's proposed Central Micronesian, or that this rule arose independently in the remaining daughter languages as a natural consequence of the fact that vowel length contrasts could not be neutralised in the environments specified by these rules.

<sup>17</sup>By no means do I intend this claim to represent an exhaustive description of PMC stress.

<sup>18</sup>It seems odd that short high vowels would devoice only after nasals. Donegan has claimed: "When voiceless vowels occur, they occur only in voiceless environments" (Donegan 1985:47). Nevertheless, this rule appears well-motivated diachronically. David Stampe (pers. comm.) has suggested that phonetic motivation for this unexpected phenomenon might be provided by assuming that, prior to (or simultaneous with) devoicing, resyllabification occurred, the result being that nasals were absorbed into preceding stressed syllables. (Therefore,  $\check{V}\$NV \text{ --->} \check{V}\$NV$ , where \$ represents syllable boundary.) Unstressed short high vowels occurring as lone segments in phrase-final syllables might then especially have been subject to lenition. Facts about MIC languages not considered here suggest that this is indeed a plausible explanation. However, if it is true that in Kiribati phrase-final short non-high vowels optionally devoice only after geminate nasals (see section 3.2), then this resyllabification hypothesis encounters difficulties. Further explanation is also required to account for the fact that Woleaian synchronically devoices phrase-final short vowels after all consonants.

### 7.3 SYNCOPE OF MEDIAL VOWELS

In addition to the preceding rules of final vowel lenition, it also seems likely that PMC had a rule of medial vowel syncope. While a complete discussion of the consonant cluster types that are synchronically permitted in MIC languages is beyond the scope of this paper, it suffices to note here that all MIC languages minimally permit geminate nasals; POC, on the other hand, exhibits no consonant clusters. The significance of this observation is that the existence of such clusters provides further evidence that the basic syllable structure of PMC had been modified so that closed syllables, at least those ending with nasals, became permissible.

To say that a language has changed its basic syllable structure from (C)V to (C)V(C), however, tells us considerably less than we need to know. Segment sequences are constrained not only syllable internally, but transsyllabically as well. Obviously, there is a substantive difference between languages permitting a wide range of consonant clusters as opposed to those like PMC, which possibly permitted only geminate nasals.

A non-linear representation of consonant clusters clearly illustrates that the transition from a strictly open syllable type to one that permitted geminate sequences represents a minimal complication of the surface segmental structure of the language. Witness the representation of the following hypothetical form:

C	V	C	V	C	V	C	V		C	V	C	C	V	C	V
											\	/			
t	a	m	i	m	a	t	a		t	a	m	a	t	a	a

Within this framework, while a change occurs on the CV tier, at the segmental level the basic alternation between consonants and vowels is maintained. Prince (1984:243) has noted: "...we expect to find many languages that allow closed syllables only when the syllable-closing consonant begins a geminate cluster; this ought to represent the next step in complication after the strict CV languages".

PMC, therefore, might have contained a constraint that permitted transsyllabic clusters within words only when they were geminate nasals. Thus, one might speculate that in PMC short high vowels were deleted between nasals word-internally and after nasals phrase-finally. In neither case were PMC transsyllabic syllable structure constraints violated. It is further likely that, as the daughter languages extended the range of possible transsyllabic consonantal sequences to geminate sonorants, homorganic nasal-obstruent sequences, and geminate obstruents, all of which are evidenced in the daughter languages exhibiting full apocope, then final vowels could be deleted in a broader variety of environments without violating any transsyllabic consonantal constraints.

### 7.4 FEATURE TRANSFER BETWEEN CONSONANTS AND VOWELS

A fourth property of PMC phonology is the apparent read-off of features between vowels and consonants. More specifically, those features relevant to vowel articulation – back, round and high – were transferred both from vowels to consonants and from consonants to vowels

Jackson (1983:309) notes that POC \**mp* merged with \**ŋp* before round vowels [+round, +back], resulting in PMC \**p'*, assumed to have been a rounded velarised bilabial stop [+high, +back, +round...]. Analogously, POC \**m* and \**ŋm* merged before round vowels, resulting in PMC \**m'*, apparently a rounded velarised bilabial nasal (Jackson 1983:312).

Marck (1975) has also argued that some vowel allophony in PMC was conditioned by the preceding consonant. PMC \**u*, he suggests, was centralised to [ʊ] ([+back ---> [-back]) when preceded by \**t*, \**d*, \**l*, \**n* and \**ŋ*, all of which he characterises as 'non-back consonants'. Synchronically, this system has expanded to play a major role in the phonological systems of Ponapean (Rehg 1981:43-46), Marshallese (Bender 1968), and possibly Kosraean. (See also Jackson 1983:323.)

#### 7.5 REGRESSIVE VOWEL ASSIMILATION

Jackson argues that all MIC languages, either synchronically or diachronically, exhibit regressive vowel assimilation – the assimilation of a vowel to a following vowel or glide. “Even in KIR, which has only five phonemic vowels and thus looks rather well-behaved, the low vowel regularly fronts in diphthongs before a front vowel, and backs before a back vowel” (1983:322). Regressive vowel assimilation is another of the thirteen phonological innovations cited by Jackson in support of reconstructing PMC.

#### 8. EROSION AND ACCRETION IN THE DAUGHTER LANGUAGES

The five phonological processes attributed to PMC – (1) penultimate stress, (2) final vowel lenition, (3) the syncope of medial vowels, (4) feature read-off between consonants and vowels, and (5) regressive vowel assimilation – represent structural properties inherited by all MIC languages. The synchronic patterns of final vowel lenition exhibited by these languages consequently may be viewed as a result of the extension or suppression of these processes. What remains to be considered is how these processes interacted so that Woleaian and Kiribati did not, like the other languages considered here, ultimately undergo full final-vowel apocope.

Sapir's view of drift is more germane here. He notes:

I would suspect...that phonetic change is composed of three basic strands: (1) A general drift in one direction, concerning the nature of which we know almost nothing but which may be suspected to be of a prevailing dynamic character (tendencies, e.g., to greater or less stress, greater or less voicing of elements); (2) A readjusting tendency which aims to preserve or restore the fundamental phonetic pattern of the language; (3) A preservative tendency which sets in when too serious morphological unsettlement is threatened by the main drift. (Sapir 1921:186)

As Sapir suggests, then, one might expect that a sound change potentially resulting in substantial loss of phonetic information would eventually be suppressed unless it developed in conjunction with a compensatory, information-preserving strategy. That is, it is reasonable to presume that in linguistic change there should be some correlation between processes of erosion and accretion.<sup>19</sup>

If one considers the five phonological processes previously attributed to PMC, one, penultimate stress, simply defines a structural property of the proto-language. Two, final vowel lenition and the

<sup>19</sup>I employ the term 'accretion' here to refer to the transfer of phonetic features from one segment to another, analogous to the way the term is employed in real estate law. (Alluvial deposits, for example, are a type of accretion.) Other information preserving strategies, of course, might also occur. Accretion could simply be additive, and potentially it could entail modification of any aspect of the grammar. It does not seem unreasonable to speculate, for example, that there might be an inverse correlation between the number of phonemes in a language and the phonemic 'length' of its morphemes, or the number of morphemes in a word.



syncope of medial vowels, are erosive. The remaining two processes, transfer between consonants and vowels and regressive vowel assimilation, are accretive. That is, these processes potentially serve to encode, at least partially, final vowel information through the creation of new consonant or vowel phonemes. The languages that (apparently) underwent full apocope – Kosraean, Marshallese, Ponapean and Lagoon Trukese – exploited one or both of the accretive processes present in PMC for just such purposes. Woleaian and Kiribati did not.

### 8.1 THE TRANSFER OF VOWEL FEATURES TO CONSONANTS

Ponapean has preserved and extended the PMC process discussed in section 7.4, whereby the high vowel /u/ became [ɯ] when preceded by 'front' consonants. Synchronically in Ponapean, all segmental phonemes belong to one of two series – front (palatal or palatalised) or back (velar or velarised). Central vowels result when front vowels occur between back consonants, or when back vowels occur between front consonants. Breaking articulation (diphthongisation) occurs when a vowel is flanked by consonants of different series.<sup>20</sup> This process operates at an allophonic level, however. It has not resulted in the addition of new segmental phonemes.

Marshallese *has* exploited this process to create new consonantal phonemes. From the five-vowel system of POC, exhibiting three degrees of vowel advancement and three degrees of height, Marshallese has developed a synchronic system in which there are four vowel phonemes, distinguished *only* in terms of height. Diachronically, the features associated with vowel advancement were transferred to the non-syllabics. Thus, in contemporary Marshallese, consonants are of one of three types – plain (or palatalised), velarised, and labialised and velarised. The vowel phonemes of Marshallese phonetically surface as front, central, or back, depending upon the nature of the adjacent consonants. For example, a high vowel between palatalised consonants surfaces as [i], between velarised consonants as [ɨ], and between labialised and velarised consonants as [u]. As in Ponapean, breaking articulation occurs when a vowel is flanked by consonants of different types.<sup>21</sup>

Kosraean is unique among MIC languages in that phonemically it exhibits three degrees of advancement in vowels and three series of consonants, phonetically parallel to those of Marshallese. While this system appears excessively complex, suggesting that significant generalisations are being overlooked, attempts to further simplify the phonemic inventory of this language (presumably in the direction of Marshallese) have thus far failed.<sup>22</sup>

At the opposite end of the spectrum from Marshallese and Kosraean are Kiribati and the Trukic languages, which exhibit three series of advancement in vowels and only a single series of consonants. The PMC process resulting in the phonetic conditioning of advancement in vowels is apparently not operative in these languages, or if any residue of such a process survives, it has not been reported.

One may thus observe that there is a balance maintained among MIC languages between complexity in consonant and vowel systems, involving four series of segments. Ponapean has two series of vowels and two series of consonants. Marshallese has a single series of vowels and three series of consonants. Kosraean alone exhibits both three series of consonants and three series of

<sup>20</sup>See Rehg (1981) for further discussion.

<sup>21</sup>See Bender (1968) for further discussion.

<sup>22</sup>See Lee (1975) for further discussion.



vowels. Kiribati, Trukese and Woleaian have maintained three degrees of advancement in vowels, with only a single consonant series. Therefore, in these three languages the transfer of vowel features to consonants noted for PMC has apparently atrophied. Nevertheless, Trukese, unlike Kiribati and Woleaian, has undergone full apocope. Consequently, one must look for further differences in the historical development of these languages.

## 8.2 REGRESSIVE VOWEL ASSIMILATION

While it seems plausible, as Jackson has argued, to assume that PMC exhibited allophonic processes of regressive vowel assimilation, it is not the case that all daughter languages subsequently exploited this process to an equal degree.

All MIC languages under consideration in this study, except precisely Kiribati and Woleaian, diachronically exhibit a stage in which a low or lower-mid back rounded vowel phoneme developed as a consequence of assimilation of /a/ to a following mid-back rounded vowel.<sup>23</sup> Trukese, Ponapean, Marshallese and apparently Kosraean further elaborated regressive vowel assimilation so that non-high vowels in medial position assimilated to word-final vowels in a manner reminiscent of Germanic umlaut. Only Kiribati and Woleaian fail to exhibit such processes. In fact, these two languages synchronically exhibit *progressive* vowel assimilation and regressive vowel *dissimilation*.<sup>24</sup>

## 8.3 SUMMARY OF EROSIVE AND ACCRETIVE PROCESSES IN MICRONESIAN LANGUAGES

The following chart summarises the observations made in the previous two sections. Languages are marked as plus for the first two processes, therefore as exhibiting these processes, only when new segmental phonemes resulted from their application.

	C ASSIMILATION TO V'S	REGRESSIVE V ASSIMILATION	FULL V APOCOPE
Kiribati	-	-	-
Marshallese	+	+	+
Kosraean	+	+	+
Ponapean	-	+	+
Lagoon Trukese	-	+	+
Woleaian	-	-	-

Just those languages that exploited accretive processes of the types described in sections 8.1 and 8.2 underwent full vowel apocope. In Woleaian and Kiribati, where these processes are/were not employed to create new segmental phonemes, thus at least in part encoding final vowel information, final vowel lenition remains highly constrained.

<sup>23</sup>The phonemic inventory of Woleaian includes a low back rounded vowel, but this vowel only occurs long; historically, it resulted from the fusion of \*eu (Sohn 1984:181). Jackson (1984:262) suggests a similar rule at an allophonic level for Proto Trukic, but, so far as I am aware, there is no evidence that such a rule ever existed in Woleaian. It can also be argued that Marshallese underwent this rule, but subsequently collapsed all phonemic contrasts of vowel advancement.

<sup>24</sup>See Bender (1973) for further discussion.

## 9. CONCLUSION

In this paper I have argued in support of the following three claims: (1) While all Micronesian languages exhibit final vowel lenition, the processes resulting in such lenition are not identical among all languages. (2) Because different processes result in parallel developments, final vowel lenition constitutes an example of drift. (3) Drift, in this instance, can plausibly be explained as a consequence of the retention, extension, or suppression of a constellation of phonological processes present in Proto Micronesian.

I have also attempted to sustain Sapir's observation that phonologically erosive processes cannot operate unchecked. Meaning can be viewed as information embedded in a sound complex, and serious diminution of that complex potentially results in loss of information. Sound change, at least long term change involving loss or neutralisation of phonemes, must then necessarily be investigated in terms of the interaction between erosion and the compensating mechanisms of accretion.

The phenomena associated with final vowel lenition in Micronesian languages, since they involve drift, provide no substantive evidence for subgrouping beyond weakly supporting the existence of a Trukic subgroup, which is strongly confirmed on other bases. Therefore, while these data do not support the more extensive subgroupings of Jackson, neither do they necessarily lead one to conclude that further subgrouping is impossible or inappropriate. Data involving drift phenomena, by their very nature, are of little or no utility in developing subgrouping arguments.

A contrapuntal consequence of this observation is that, in developing subgrouping hypotheses, one must take into account the possibility of drift. Drift, as I hope to have demonstrated in this paper, does indeed represent a natural type of linguistic evolution.

## ACKNOWLEDGEMENTS

I dedicate this paper to George Grace, with affection and admiration.

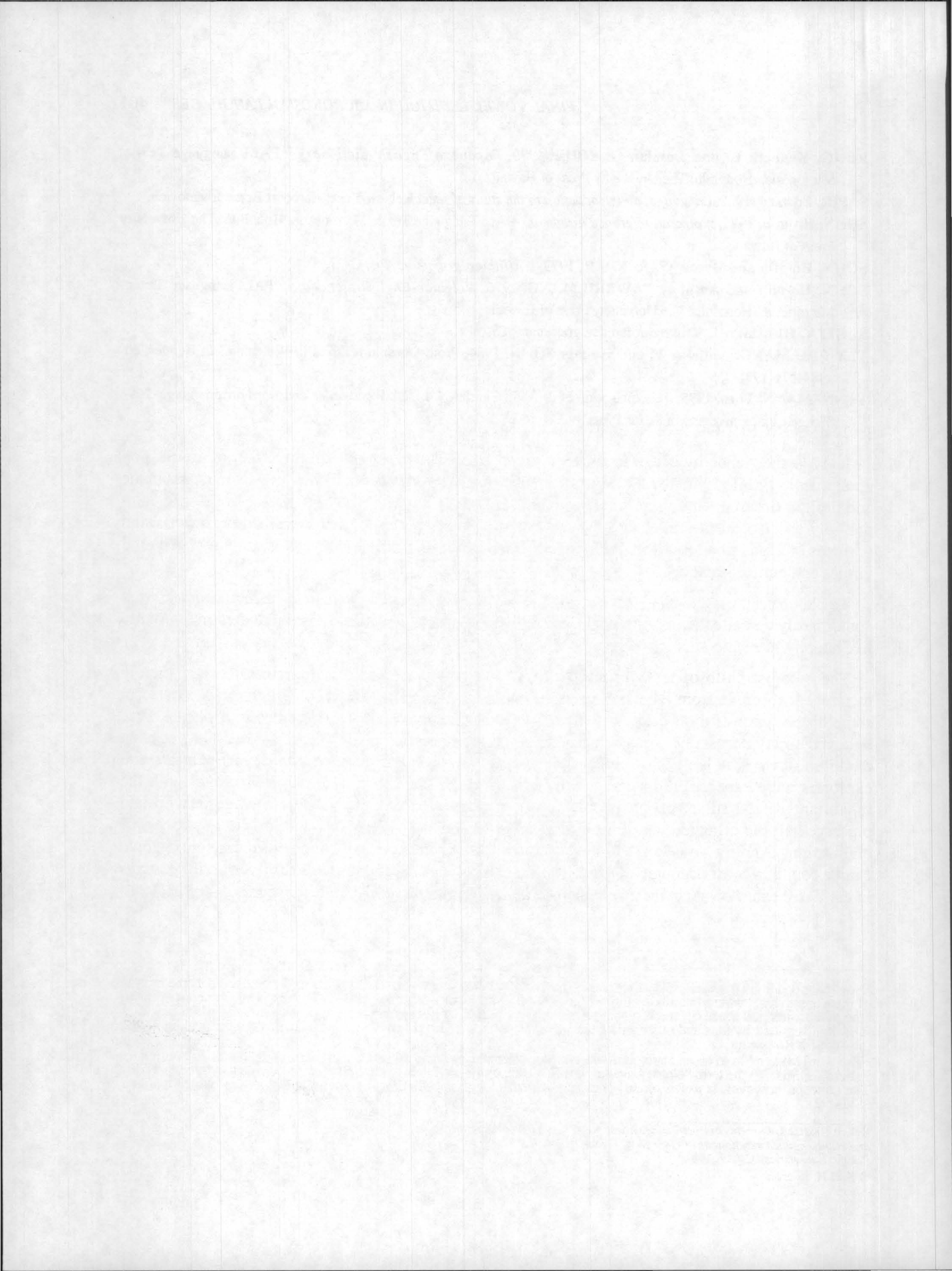
I wish to thank Byron Bender, Robert Blust, Iovanna Condax, Patricia Donegan, and Ann Peters for their helpful comments on an earlier draft of this paper.

## BIBLIOGRAPHY

- ABO, Takaji, Byron W. BENDER, Alfred CAPELLE and Tony DeBRUM, 1976, *Marshallese-English dictionary*. Honolulu: University Press of Hawaii.
- BENDER, Byron W., 1968, Marshallese phonology. *Oceanic Linguistics* 7/1: 16-35.
- 1969, The other orthographic principle. Talk given to May 6 meeting of the Linguistic Society of Hawaii. TS.
- 1971, Micronesian languages. In Thomas A. Sebeok, ed., *Current trends in linguistics*, vol.8, part i: *Linguistics in Oceania*, 426-465. The Hague: Mouton.
- 1973, Parallelisms in the morphophonemics of several Micronesian languages. *Oceanic Linguistics* 12 (1 & 2): 455-477.
- BENDER, Byron W. ed., 1984, *Studies in Micronesian linguistics*. PL, C-80.
- BENDER, Byron W., Robert W. HSU, Frederick JACKSON, Kenneth L. REHG, Stephen TRUSSELL and Judith W. WANG, 1983, Micronesian cognate sets. Computer printout. Honolulu: Department of Linguistics, University of Hawaii.
- BENDER, Bryon 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 (Memoir 43)*. Wellington: The Polynesian Society.
- 1990, Three recurrent changes in Oceanic languages. In J.H.C.S. Davidson, ed. *Pacific island languages: essays in honour of G.B. Milner*, 7-29. London: School of Oriental and African Studies, University of London; Honolulu: University of Hawai'i Press.
- BYNON, Theodora**, 1977, *Historical linguistics*. Cambridge University Press.
- DONEGAN, Patricia J.**, 1985, *On the natural phonology of vowels*. Outstanding Dissertations in Linguistics. New York: Garland.
- DONEGAN, Patricia Jane and David STAMPE**, 1983, Rhythm and the holistic organization of language. In J.F. Richardson, et al., eds *Papers from the parasession on the interplay of phonology, morphology and syntax*, 337-353. Chicago: Chicago Linguistic Society.
- DYEN, Isidore**, 1949, On the history of Trukese vowels. *Language* 25:420-436.
- 1965, *A sketch of Trukese grammar*. American Oriental Series, Essay 4. New Haven: American Oriental Society.
- ELBERT, Samuel H.**, 1974, *Puluwat grammar*. PL, B-29.
- FOLEY, James**, 1977, *Foundations of theoretical phonology*. Cambridge University Press.
- GOODENOUGH, Ward H. and Hiroshi SUGITA**, 1980, *Trukese-English dictionary*. Philadelphia: American Philosophical Society.
- 1990, *Trukese-English dictionary supplementary volume: English-Trukese and index of Trukese word roots*. Philadelphia: American Philosophical Society.
- GRACE, George W.**, 1969, A Proto-Oceanic findex. *Working Papers in Linguistics, University of Hawaii* 1/2:39-84.
- GROVES, Tereb'ata R., Gordon W. GROVES and Roderick JACOBS**, 1985, *Kiribatese: an outline description*. PL, D-64.
- HARRISON, Sheldon P.**, n.d., Kiribati reference grammar. TS.
- JACKSON, Frederick H.**, 1983, The internal and external relationships of the Trukic languages of Micronesia. PhD dissertation, University of Hawaii.
- 1984, Reflexes of Proto-Oceanic in the Trukic languages of Micronesia. In Bender, ed. 1984:259-280.
- LAKOFF, Robin**, 1972, Another look at drift. In Robert P. Stockwell and Ronald K.S. Macaulay, eds *Linguistic change and generative theory*, Bloomington: Indiana University Press.
- LEE, Kee-dong**, 1975, *Kusaiean reference grammar*. PALI Language Texts: Micronesia. Honolulu: University Press of Hawaii.
- 1976, *Kusaiean-English dictionary*. PALI Language Texts: Micronesia. Honolulu: University Press of Hawaii.
- LEE, Kee-dong and Judith W. WANG**, 1984, Kosraean reflexes of Proto-Oceanic phonemes. In Bender, ed. 1984:403-422.
- MARCK, Jeffrey C.**, 1975, The origin and dispersal of the Proto-Nuclear Micronesians. MA thesis, University of Iowa.
- 1977, A preliminary manual of Micronesian sound correspondences. Paper presented at the Symposium on Austronesian Linguistics at the Summer Meeting of the Linguistic Society of America. Honolulu, Hawaii. MS.
- PRINCE, Alan S.**, 1984, Phonology with tiers. In Mark Aronoff and Richard T. Oehrle, eds *Language sound structure*, 234-244. Cambridge: The MIT Press.
- REHG, Kenneth L.**, 1981, *Ponapean reference grammar*. PALI Language Texts: Micronesia. Honolulu: University Press of Hawaii.
- 1984a, On the history of Ponapean phonology. In Bender, ed. 1984:281-316.
- 1984b, The origins of 'compensatory lengthening' rules in Micronesian languages. In Bender, ed. 1984:53-59.
- 1986, The phonology of Ponapean vowels. PhD dissertation, University of Hawaii.
- REHG, Kenneth L. and Byron W. BENDER**, 1990, Lexical transfer from Marshallese to Mokilese: a case of intra-Micronesian borrowing. *Oceanic Linguistics* 29/1: 1-26.

- REHG, Kenneth L. and Damian G. SOHL, 1979, *Ponapean-English dictionary*. PALI Language Texts: Micronesia. Honolulu: The University Press of Hawaii.
- SAPIR, Edward, 1921, *Language: an introduction to the study of speech*. New York: Harcourt Brace Jovanovich.
- SOHN, Ho-min, 1975, *Woleaian reference grammar*. PALI Language Texts: Micronesia. Honolulu: The 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*. PALI Language Texts: Micronesia. Honolulu: The University Press of Hawaii.
- SUGITA, Hiroshi, n.d., *Trukese reference grammar*. TS.
- TAWERILMANG, Anthony F. and Ho-min SOHN, 1984, Proto-Oceanic reflexes in Woleaian. In Bender, ed. 1984:171-199.
- VENNEMANN, Theo, 1975, An explanation of drift. In Charles N. Li, ed. *Word order and word order change*, 269-305. Austin: University of Texas Press.



THE LANGUAGE OF THE NOBLE SAVAGE:  
EARLY EUROPEAN PERCEPTIONS OF TAHITIAN

KARL H. RENSCH

1. THE CONCEPT OF THE NOBLE SAVAGE

In 1749 the Academy of Dijon awarded the prize for the best essay on the effect of civilisation upon man to a philosopher by the name of Jean-Jacques Rousseau who had expressed in a passionate and defiant diatribe what quite a few intellectuals at the time felt, but were reluctant to proclaim publicly as a social doctrine, that the development of civilisation, and the advances in arts and sciences had led to a loss of naturalness in human behaviour and contributed to a decline of moral virtues and ethical standards.

To blame progress and change for the demise of mankind and to indulge in a *laudatio temporis acti* could hardly be called an original idea. The lament about paradise lost where innocent man used to live happily in a state of unspoiled nature is a recurrent theme throughout the history of mankind.

The escapist philosophers of the eighteenth century conceptualised paradise not as a merely metaphysical refuge from the harsh realities of life but as geographical loci which really existed in some distant parts of the world<sup>1</sup>. Atlantis, the Hesperides and Avalon, blessed isles of plenty, were believed to lay across the western oceans. As some remote areas of the globe had not yet been explored, there was hope, that an earthly paradise would be discovered one day. Seafarers and explorers were expected to bring back the news and tell the world about their encounter with the inhabitants of paradise, happy people living a simple, natural life, Noble Savages,<sup>2</sup> who had preserved all the original qualities, physical and moral, that modern man had lost in the process of civilisation. And the voyagers willingly obliged. Romanticising glorifications of indigenous people can be found in most accounts of voyages to faraway lands. Columbus's description of the Indians in the Carribean shows that they were already à la mode during the age of discovery.

---

<sup>1</sup>Lord Monboddo, in a serious scientific work, asserted in so many words, that the Golden Age yet lingered in the islands of the South Seas, "where the inhabitants live without toil or labour upon the bounty of Nature". In Otaheite, he says, "the inhabitants pull bread off trees, which grow with no culture, for about nine months of the year, and when this food fails, it is supplied by nuts and other wild fruits" (Tinker 1964:11). Lord Monboddo (James Burnett) became known as the Scottish Rousseau.

<sup>2</sup>The word savage has to be understood in its etymological sense as *homo silvaticus* = man from the bush. According to Fairchild (1961:29) the term "Noble Savage" was first used by Dryden in his play *Conquest of Granada*, written in 1701. The concept, however, is much older. It can be traced back to speculations by social philosophers like Montaigne (1894:92).



In the eighteenth century the enthusiastic reports on New Cythera (Tahiti) published by members of Bougainville's exploring expedition around the world convinced many contemporaries that paradise had at long last been found. The description of the inhabitants of Tahiti miraculously fitted Rousseau's vision of the Noble Savage as the *alter homo*, the opposite of modern man. 'Otaheitans' were depicted as free from malice and vanity, with no interest in the accumulation of wealth and power. Content on satisfying their physical needs by accepting what nature had to offer, they enjoyed sensual pleasures unconstrained by the kind of repressive social customs that existed in Europe. And as far as their language was concerned, unlike the corrupted tongues of the Occident, it was melodious, expressive and natural.

## 2. THE LANGUAGE OF THE NOBLE SAVAGE; SOCIOPHILOSOPHICAL SPECULATIONS

To get an idea to what extent the minds of scientists and intellectuals had been influenced by the sociophilosophical discussions in the wake of Rousseau's theories, one only has to look at Commerson's description of the Tahitian language. Philibert de Commerson, physician and botanist, accompanied Bougainville as a naturalist on his expedition. In a letter written in 1769 at Isle de France (Mauritius) and published in the *Mercure de France* in November of the same year, he talks about his encounter with Tahiti, the island he calls 'Utopia' or 'Fortunate' and the idiom of its inhabitants:

A very sonorous, very harmonious language, composed of about four or five hundred words lacking in declension or conjugation, that is completely without syntax, is sufficient for them to render all their ideas and to express their every need. It is characterized by a noble simplicity which, excluding neither tonal modifications nor emotional pantomime, protects it from the arrogant tautology which we call richness of language, which make us lose niceness of perception and speed of judgement in a maze of words. The Tahitian, to the contrary, names his object as soon as he perceives it. The tone in which he has pronounced the name of the object has already indicated how he is affected by it. Paucity of words makes for rapid conversation. The operation of the soul, the movements of the heart, are isochronous with the movements of the lips. The speaker and the listener are always in unison. Our Tahitian prince, who in the seven or eight months that he had been with us, had not yet learned ten of our words, most often stunned by their volubility, had no other expedient than to stop his ears, and laugh in our faces. (Knowlton 1955:8)

Commerson's remarks on the Tahitian language are representative for many of the reports that we have of the early visitors to the Society Islands. They strike us by their charming simpleness in presenting personal interpretations as objective descriptions of facts, which invariably give credence and support to the writer's philosophical stance or view of the world. Reading Commerson's description of Tahitian one has little difficulty in identifying him as a fervent believer in Rousseau's theories on the origin and role of human language as expounded in his essay 'On the inequality of man'.

Language, according to Rousseau, started with the "cri de nature", the urge to express inner feelings of love, hate, compassion and anger by unarticulated cries.<sup>3</sup> The close link between vocalisation of feelings and the shape of words is still preserved in the language of those people who

<sup>3</sup>"Ce n'est ni la faim ni la soif, mais l'amour, la haine, la pitié, la colère, qui leur ont arraché les premières voix" (Rousseau 1974:96).

live in a state of nature. Their language is characterised by the prevalence of vowels over consonants. To make up for the phonotactic drawback, that is, the limited variability of syllables due to the small number of consonants (which are seen as conventional devices, *thesei* rather than *physei*), the languages resort to natural means to make up words without creating homonyms: increasing the number of vowel quality distinctions, using accents, adding quantity and varying the rhythm. As a result, natural languages are melodious and harmonious showing a close affinity with poetics and music.

When Commerson emphasises that the Tahitian language is melodious and harmonious, he does so not only in order to praise its aesthetic quality. He is also anxious to portray Tahitian as a language which is complete in the sense that Rousseau (1974:145) had talked about this aspect with reference to the Greek language:

Une langue qui n'a que des articulations et des voix n'a donc que la moitié de sa richesse: elle rend des idées, il est vrai; mais pour rendre des sentiments, des images, il lui faut encore un rythme et des sons, c'est-à-dire, une mélodie; voilà ce qu'avait la langue grecque, et ce qui manque à la nôtre.

The Tahitian language has more in common with Greek than with French. Its sound inventory may lack the consonants that Greek required for building up its extensive lexicon, but as the communicative needs of people living in a natural environment are less demanding – Commerson estimates the vocabulary to consist of 500 words only – Tahitian cannot be called inferior just on account of its limited lexicon. What is important apart from being capable of adequately rendering the ideas and notions which the language user considers relevant, is the capacity of Tahitian to express feelings and emotions. Commerson has this particular aspect in mind when he twice makes reference to the use of tone in Tahitian: “It is characterized by a noble simplicity which, excluding neither tonal modifications nor pantomime”, and later on “the tone in which he has pronounced the name of the object has already indicated how he is affected by it” (Knowlton 1955:10).

While the naturalness of the Tahitian language manifests itself through sonority and tonality, its simplicity, another fundamental virtue of the Noble Savage's language, is guaranteed by the lack of declensions and conjugations and the complete absence of syntax. The advantages are self-evident. Unlike French, the Tahitian language will not confuse the listener by what Commerson calls an “arrogant tautology” resulting from complicated grammatical structures and a maze of words. Although he doesn't mention the term ‘classical simplicity’, Commerson is obviously guided by this notion which in European literary tradition is closely linked to Greek antiquity. The Greek and the Tahitian language may differ in internal structure and external appearance, but they share the fundamental quality of being classical in their completeness and simplicity. Tahiti is the New Cythera also on account of its language.

Further characteristics of the language of the Noble Savage are its honesty and its potential for spontaneous creativity. Commerson: “the operations of the soul, the movement of the heart are isochronous with the movement of the lips” and “the Tahitian names his object as soon as he perceives it” (Knowlton 1955:10). While ‘naturalness’ and ‘simplicity’ are qualities for which Commerson cited melody/harmony and simple grammar/small lexicon as evidence, these characteristics are of a less tangible nature. They only surface when communication problems have to be dealt with. Then, because there is a direct relationship between feeling and expression, the native speaker will show his ability to spontaneously verbalise new concepts and cope with any lexical

requirement that the new situation brings about. In other words, the Tahitian language has an inbuilt potential for growth and expansion.

### 3. FIELDWORK IN SITU; TAHITIAN UNDER CLOSE SCRUTINY.

Commerson's remarks on the Tahitian language belong to the realm of philosophy of language. They tell us very little about the linguistic properties of the language itself. His primary concern was to convince his contemporaries that Tahitian had exactly those qualities they had expected of the language of the Noble Savage. While Bougainville and his companions Duclos-Guyot and Fesche compiled word lists of the language of New Cythera (Fesche befittingly using Greek characters), the first attempts to analyse Tahitian, especially its sound system, were made by members of Cook's first voyage to the Pacific. Not concerned about presenting the natives as the incarnation of Rousseau's idea of man's superiority in his natural state, they had no qualms about making negative remarks on aspects of the Tahitian language which they considered inadequate, or quoting examples of the inability of Tahitians to properly pronounce English words. Banks writes in his *Endeavour journal*:

Their Language appeard [sic] to me to be very soft and tuneable. It abounds much with vowels and was very easily pronounc'd by us when ours was to them absolutely impracticable. I shall instance particularly my own name which I took much pains to teach them and they to learn: after three days fruitless trials I was forc'd to select from their many attempts the word *Tabáne*, the only one I had been able to get from them that had the least similitude to it. Again Spanish or Italian words they pronouncd [sic] with ease provided they ended with a vowel, for few or none of theirs end with a consonant...In one respect however it is beyond measure inferior to all European languages, which is its almost total want of inflexion both of Nouns and verbs, few or none of the former having more than one Case or the latter one tense. (Beaglehole, ed. 1959, I:370)

Parkinson's (1773:65) description of the difficulties that Tahitians had with the pronunciation of English is even more detailed:

The natives could not repeat after us the sound of the letters, Q, X, and Z, without great difficulty; G, K, and S, they could not pronounce at all. Many of the names of the people of our ship having the G, K, or S, in them, they could not approach nearer the sound of them as follows:

Toote	Cook	Mata	Monkhouse
Opane	Banks	Petrodero	Pickersgill
Tolano	Solander	Tate	Clark
Treene	Green	Poline	Spoving
Hite	Hicks	Taibe	Stainsby
Towara	Gore	Patine	Parkinson

Parkinson's reference to the letters *q*, *x* and *z* is somewhat ambiguous as we don't know whether they were presented to the Tahitians as individual letters pronounced [kju], [eks] and [zet] or in the context of words like quick, box or zero. In any case it is nearer to the truth that the difficulties were such that the Tahitians couldn't pronounce them at all as they contain a velar stop and/or a sibilant. As the examples (Solander > *Tolane*, Cook > *Toote*, Hicks [hiks] > *Hite*) show both *s* and *k* were replaced by an alveolar stop. The transliteration *Opane* for Banks is to be read as *o* (personal article

as in *Omai*) + *pane* thus confirming the *-bane* fragment in *Tabane* (*ta* = definite article *te?*) in the version quoted by Banks himself. For *Gore* one would have expected *Tore* instead of *Towara*. The simplification of consonant clusters is quite radical, but the strategies seem to be less rule governed: *nks* > *n* (Banks), *nk* > *t* (Monkhouse) *sg* > *d* (Pickersgill), *kl* > *t* (Clark), *nsb* > *b* Stainsby, *ns* > *n* (Parkinson), *ks* > *t* (Hicks), *sp* > *p* (Spoving), *st* > *t* (Stainsby). In *Treene* and in *Petrodero* *tr* was probably pronounced with a transitional unstressed schwa which in rapid speech is hardly audible especially when the preceding or following syllable is stressed.

Parkinson does not provide an explanation for the inability of the Tahitians to produce certain consonants and consonant clusters. He contents himself with observing the facts without inferring lack of intelligence or cognitive ability. After all, the English have similar problems with some of the Tahitian sounds:

They have various sounds peculiar to themselves, which none of us could imitate; some of them they pronounced like B and L mingled together; other between B and P, and T and D. Some like Bh, Lh, and Dh. (1773:66)

Parkinson's remark on *b/p* and *t/d* is reflected in the transliteration of Banks as *Opane/Tabane*, of Pickersgill as *Petrodero* and in the inconsistent use of these letters in Tahitian word lists compiled by various members of the expedition. The English had difficulties in deciding on the nature of the bilabial and alveolar stops in Tahitian. In Polynesian languages, there is no phonological distinction between voiced and unvoiced stops. Furthermore, vowels following a stop are never realised with delayed voice onset, in other words, stops are not aspirated. In English *b* and *p* and *d* and *t* are phonemically distinct, but in prevocalic position, the voiceless stop is also aspirated if it is not part of a cluster. For phonetically untrained native speakers of English, it is the presence of aspiration and not so much the absence of voice that provides the main clue for the identification of *p* and *t* (and *k*). Unaspirated *p* and *t* sound to an English ear very much like *b* and *d*, but in the case of Tahitian the confusion was further aggravated on the production side by the absence of a voiced-unvoiced contrast which allowed the speaker to vary the realisation of bilabial and alveolar stops without transgressing into the allophonic domain of another phoneme sharing the same point of articulation.

#### 4. IS THE NOBLE SAVAGE LINGUISTICALLY DEFICIENT? TAHITIANS PUT TO THE SHIBBOLETH TEST

Banks's and Parkinson's observations of Tahitian pronunciation, while unbiased and rather specific as compared to Commerson's general and uncritical remarks, were not based on a systematic study and testing of informants. They were meant to be examples of the distinct nature of the Tahitian and English sound inventory and the articulatory constraints which this difference imposed on those who try to speak the other's language. Completeness was never intended and the authors never claimed to be experts in foreign languages, let alone phonetics. The first 'scientific' study of Tahitian by a linguist of sorts using a native speaker as an informant was not carried out in the field, but in a studio in Paris on 25 April 1769. On this day Jacob Pereire, a Spaniard who had invented a method of teaching the deaf and dumb, met Ahutoru<sup>4</sup> from Tahiti who had been brought to France by Bougainville.

The arrival of the first Polynesian in Europe in March 1769 had caused a sensation. Here at last was living proof of the existence of the Noble Savage. Before him, exotic visitors from faraway

<sup>4</sup>His name is traditionally, although incorrectly spelled Aoutourou or Aotourou.

lands had been transported to the old continent to be scrutinised by erudite armchair philosophers and self-appointed anthropologists: Eskimos from Labrador, a Brazilian 'king', Indians from Guiana, a 'genuine' princess from Virginia. None of them had lived up to the expectation of academia and the salons. They became objects of curiosity for the general public at fairs and circus performances like Epenow, the giant redskin who was "shown up and down London for a wonder" (Alexander 1977:69).

Ahutoru and the other two visitors from the Pacific, Ma'i from Huahine and Lee Boo from Belau who arrived later escaped the fate that their predecessors experienced in Europe. They turned up when the stage was set. They were going to fill a role at an historical moment when actors were needed to play it.<sup>5</sup> And they played that role admirably well. There was general agreement in Paris and London that the social behaviour of the visitors was characterised by naturalness, sincerity and gracefulness which put the artificiality of western etiquette to shame. Befriended and protected by influential men, they had access to the highest circles of society. The disillusionment that had set in with other visitors as soon as they started interacting in their new environment never happened. As far as their private and public conduct was concerned they were almost beyond criticism as their idiosyncracies were interpreted as a reaction against the unnaturalness of rules and regulations dominating civilised man's life. There was, however, one issue which was disturbing and puzzling. It was very obvious to everyone who met Ahutoru, Ma'i or Lee Boo that they performed miserably when it came to speaking the language of their hosts. Was the Noble Savage intellectually inferior after all?

Commerson, as one would expect, rejects any explanation that would tarnish the image of the Noble Savage. For him the fault doesn't lie with the language learner, but with the language he is trying to learn. It is the 'volubility' of French that leaves Ahutoru stunned and demotivates him to such a degree that in a gesture of helpless frustration he stops his ears and laughs into the faces of his interlocutors. Bougainville, who had already become aware of his guest's difficulties with French on the return journey, offers a similar explanation. An entry in his diary made on 11 May, only a few weeks after leaving Tahiti, reads: "Our language is almost impossible to pronounce for somebody whose language consists only of vowels. The very same reasons which are cited to describe French as a language that is not very musical, make it inaccessible to his speech organs." But he adds that this may not be the only explanation. Ahutoru's laziness could be another factor preventing him from making satisfactory progress in French.

The examination of Ahutoru by the linguistic expert Monsieur Pereire was supposed to clarify what caused his problems with the French language. Two possibilities had to be considered: (a) Ahutoru might be of very limited intelligence (a finding that could challenge Rousseau's doctrine of the superior qualities of the Noble Savage), or (b) Ahutoru might be suffering from a physical impediment of his speech organs which prevented him from performing certain articulatory tasks (the Noble Savage was excused). The meeting was arranged by the famous mathematician and astronomer Charles de la Condamine. M. Pereire kept notes of this and all subsequent sessions and sent a transcript of his findings to Bougainville.

De la Condamine first administered a kind of IQ test.<sup>6</sup> Not unexpectedly Ahutoru scored to the tester's satisfaction. Possibility (a) being eliminated, Pereire took over. The linguistic expert started

<sup>5</sup>This is a modification of Alexander's statement which only refers to Ma'i: "The stage was set: Omai was to fill a role at a historical moment when an actor was needed to play it..." (Alexander 1977:70).

<sup>6</sup>"La Condamine décide d'examiner son intelligence par des méthodes qui ne sont pas d'une rigueur scientifique absolue: il lui fait manier un cornet acoustique" (Mazellier, ed. 1978:191).



off by asking Ahutoru to repeat after him the sounds of the French language, one by one. Ahutoru tried very hard, but to the amazement of the interviewer, was absolutely incapable of pronouncing the consonants in the syllables *ca, da, fa, ga, sa, za*, the palatal *l* or any of the nasalised vowels. But that was not all. Ahutoru could not make a distinction between *cha* and *ja*, pronounced *b* and *l* *imparfaitement* and even more so the initial *r*. Parisian *r* is a uvular fricative and the closest that Ahutoru would have ever gotten to it was probably an apical trill which exists in Tahitian. Pereire deserves credit for the accuracy of his auditory discrimination. When he asked Ahutoru to pronounce initial *r* in the Tahitian words *rai* 'big' and *roa* 'long' he noticed that it was preceded by a vowel "...car j'ay essayé de luy faire prononcer l' r à plusieurs reprises et il ne la prononçoit que très imparfaitement en traissant plustot la voyelle qui précédoit l' r" (Taillemite 1977, I:488).

Ahutoru had said *e rai* and *e roa* using the verbal particle *e* in order to produce a grammatically correct utterance, as *rai* and *roa* never occur by themselves. Pereire was wrong to conclude that *r*-initial words do not exist in Tahitian and that the words *rai* and *roa* recorded in Bougainville's word list contain spelling errors, but the logic of his argument is flawless. If Tahitians only have an intervocalic *r*, then they would find it difficult to pronounce a French word that has initial *r*, just like French people, whose language has a palatal *l* intervocalically, find it hard to pronounce this sound in word-initial position in Spanish.

After establishing which sounds Ahutoru could not produce, Pereire examined the word list that Bougainville had compiled, pointing out some dubious transcriptions like *abobo* 'demain', *allelo* 'langue', *maglli* 'froid' and the phonotactically impossible *taomta* 'couverture de tête' (an obvious printing error for *taumata*).

He criticised the use of the letter *b* as one might be led to believe that it represents "...l'articulation franche du *b*, lettre que pourtant il ne prononce qu' à l'espagnole ou sans presque joindre les lèvres" (Taillemite 1977, I:488). Spanish /b/ has two allophones: one is realised as a bilabial voiced fricative [β] occurring between vowels, and the other is realised as a bilabial stop [b]. Pereire, describing Ahutoru's pronunciation, was referring to the former when he said that the lips hardly come together in the production of this sound. It is difficult to believe that the Tahitian *p* was realised as a fricative; what he had observed was probably a *b* pronounced with reduced articulatory force, in other words, a lenis. The results of Pereire's investigations are mentioned in Bougainville's *Voyage autour du monde*. Johann Reinhold Forster, who translated the book into English, added a footnote to the passage where Bougainville (1772:272) says:

The language of Taiti is soft, harmonious and easy to be pronounced; its words are composed of almost mere vowels, without aspirates.

Forster's footnote reads:

The contrary, of the last mentioned circumstance, has been observed by our English navigators; and it is therefore highly probable Mr. de B. picked up his vocabulary of words from Aotourou, who had an impediment in his speech.

Forster had obviously misunderstood and misinterpreted what Pereire had said about Ahutoru's capacity to learn French, or rather he was led astray by the manner in which Bougainville (p.272) had presented the findings: "M. Pereire, celebrated for his art of teaching people, who are born deaf and dumb, to speak and articulate words, has examined Aotourou several times, and has found that he could not naturally pronounce most of our consonants, nor any of our nasal vowels".



Not only did Forster wrongly conclude that Ahutoru had a speech impediment, he also expressed grave doubts about his intelligence. The following statement (p.265), entirely based on hearsay, makes the Noble Savage appear in a very unfavourable light:

Though our author has strongly pleaded in this paragraph in behalf of Aotourou, it cannot, however, be denied that he was one of the most stupid fellows; which not only has been found by Englishmen who saw him at Paris, during his stay there, and whose testimony would be decisive with the public, were I at liberty to name them; but the very countrymen of Aotourou were, without exception, all of the same opinion, that he had very moderate parts, if any at all.

When Forster translated Bougainville into English, he had not yet been to Tahiti and he had never met a native speaker of the Tahitian language. His remarks are as biased as Commerson's uncritical laudation, leaving the reader in no doubt about his stance in the Noble Savage controversy. Bougainville's final assessment (p.264) which took into account Pereire's findings is far more balanced and would satisfy most psycholinguists and experts in foreign language learning of today:

Some other sharp critics conceived and propagated a very mean idea of the poor islander, because after a stay of two years amongst Frenchmen, he could hardly speak a few words of the language. Do not we see every day, said they, that the Italians, English, and Germans learn the French in so short a time as one year in Paris? I could have answered them perhaps with some reason, that, besides the physical obstacle in the organs of speech of this islander, (which shall be mentioned in the sequel) which prevented his becoming conversant in our language, he was at least thirty years old; that memory had never been exercised by any kind of study, nor had his mind ever been at work; that indeed an Italian, an Englishman, a German could in a year's time speak a French jargon tolerably well, but that was not strange at all, as these strangers had a grammar like ours, as their moral physical, political, and social ideas were the same with ours, and all expressed by certain words in their language as they are in French; that they had accordingly no more than a translation to fix in their memory, which had been exerted from their very infancy. The Taiti-man, on the contrary, only having a small number of ideas, relative on the one hand to a most simple and most limited society, and on the other, to wants which are reduced to the smallest number possible; he would have been obliged, first of all, as I may say so, to create a world of previous ideas, in a mind which is indolent as his body, before he could come so far as to adapt to them the words in our language, by which they are expressed.

If we were to rephrase this statement using the terminology of modern theories of foreign language acquisition, we could say that Bougainville had correctly identified some fundamental factors influencing the learner's proficiency, for example maturational constraints, phonological and lexical interference from the first language and sociopragmatic incongruence of Tahitian and French. And of course he was absolutely right in his reply to those who were criticising Ahutoru, that his poor performance in French was no reflection on his intelligence and could not be used as an argument in the discussion whether man in his natural state was inferior or superior to man civilised.

## 5. MA'I'S LOW ENGLISH PROFICIENCY DOES NOT DEMYSTIFY THE NOBLE SAVAGE

Ma'i (elsewhere Omai, Omy, Omiah) was an immediate success in English high society when he arrived in England on the *Adventure* in 1773. Fanny Burney commented: "He appears in a new world like a man who had all his life studied the Graces, and attended with unremitting application and diligence to form his manners, and to render his appearance and behaviour politely easy, and thoroughly well bred!" Comparing him with Mr. Stanhope, Lord Chesterton's son, who, despite his father's educational efforts, had turned out "a meer *pedantic booby*", she concludes "this shows how much more *nature* can do without *art*, than art with all her refinement unassisted by *nature*" (Alexander 1977:91). Ma'i also gave a splendid demonstration of his "unspoiled intellectual powers" by beating Signor Baretta at chess. Even Dr Johnson, one of the most outspoken critics of Rousseau, grudgingly acknowledged his potential. But as with Ahutoru in Paris, Ma'i's poor English proficiency nonplussed his admirers. How could it be that the Noble Savage performed so miserably in the language learning task?

Numerous anecdotes were in circulation, which described the humorous aspect of his attempts to master the phonetics of English and to come to grips with its syntax. But whether it was Anglo-Saxon respect for a person's dignity or a lack of Cartesian rigour, Ma'i was never subjected to the kind of linguistic test that Ahutoru had to endure. We are therefore left with rather anecdotal and sometimes even contradictory evidence such as the following where according to one source, he could not produce *g* [dʒ], having greeted His Majesty King George III with "How do, King Tosh." Fanny Burney, however, claimed that "he can pronounce the *th* as in *thank you*, and the *w* as in *well*, and yet cannot say *g* which he uses a *d* for. But I now recollect, that in the beginning of a word, as *George*, he *can* pronounce it" (Alexander 1977:89).

George Colman, who was a little boy when he became friends with Ma'i, reports in *Random Collections* that they had made up a lingo for themselves, half Otaheitan, half English, in which they contrived to jabber to their mutual enlightenment (Tinker 1964:79). Fanny Burney, whose brother Lieutenant Jem Burney had been to Tahiti with Furneaux, even supplies us with bits and pieces of discourse in pidgin à la Omai.

He began immediately to talk to me of my brother.

"Lord Sandwich write one, two, three" (counting on his fingers) "*monts* ago,— Mr Burney — come home."

"He will be very happy," cried I, "to see *you*."

He bowed and said, "Mr Burney very *dood* man!"

We asked if he had seen the King lately?

"Yes; King George *bid me*, — 'Omy, you go home.' Oh, very *dood* man, King George!" (Alexander 1977:135)

Although Ma'i had great difficulties with English, he was an excellent communicator who got the message across using gestures and body language when his linguistic resources were exhausted. As far as the Rousseauists were concerned, his communicative skills compensated for his low proficiency in speaking English correctly as it manifested the natural intelligence, ingenuity, inventiveness and resourcefulness expected of a Noble Savage. The man from Raiatea very cleverly transferred the conceptual categories that he was familiar with to the new sociocultural context by producing interlanguage words and phrasal expressions which impressed and amused the English public. These creations, known as Omaisms, made the round of the salons.

Ma'i, who had never seen ice, called it "stone-water", and for snow which was equally unknown to him, he came up with the term "white rain". At a reception given by Lord and Lady Sandwich at Hinchinbrook, he was asked what he liked best about London:

"The great hog that carries people" he replied instantly. "English hogs ver'fine", he continued. "Only this morning Lord an Lady Sandwich show me the great hog that gives coconut milk. Ver' good. No climb tree - only put hand under hog and squeeze." (Clark 1969:26)

When offered a pinch of snuff, Ma'i gracefully declined, saying:

"No tank you, Sir. Me nose be no hungry!" He called Captain Furneaux "King of the Ship", Lord Admiral Sandwich "King of all the Ships" and the butler "King of the Bottles". (Clark 1969:27)

Cradock recalled another incident at Hinchinbrook which had people admire his linguistic ingenuity:

He came in whilst we were at breakfast at Hinchinbrook, his hand was violently swelled, and he appeared to be in great agony, but could not explain the cause. At last, not being in possession of the word wasp, he made us understand that he had been wounded by a soldier bird. We were all astonished: and Dr. Solander very well remarked, that considering the allusion to the wings and the weapon, he did not know that any of the naturalists could have given a more excellent definition. (Alexander 1977:83)

According to Fanny Burney Ma'i's English became better as time went by. Twelve months after she had first met him she wrote:

Since his first visit, twelve months before he had learned a great deal of English and, with the aid of signs and actions, could make himself tolerably well understood. He pronounced the language differently from other foreigners, sometimes unintelligibly, but he had really made great proficiency, considering the disadvantages he laboured under; for he knew nothing of letters, while so very few persons were acquainted with his tongue that it must have been extremely difficult to instruct him at all...our conversation was necessarily very much confined; indeed it wholly consisted in questions of what he had seen here, which he answered, when he understood, very *entertainingly*. (McCormick 1977:157)

Ma'i had started to learn English aboard the *Adventure* on the voyage to England from Jem Burney, who himself had a keen interest in the Tahitian language. Ma'i was highly motivated and worked very hard. He used a rather unusual mnemonic method to memorise English words. Unable to read and write, he tied a knot into a cord for each new lexical item that he had learned<sup>7</sup>. The knot-association device helped him to visualise and recall them, he had in a manner of speaking the newly learned words at his fingers' end.

Tane, man; wahine, woman; upoo, the head; niho, the teeth; iore, rat; wai, water; vaa, canoe; tira, mast; inu, sleep(sic!). These words and many others, he repeated to himself as he told off his rosary of knowledge. On it went knots representing parts of the body, table utensils, ship's articles, colors, numerals, and the name of every ship they met at sea

<sup>7</sup>According to a missionary source a 'writing system' using knots was used by the inhabitants of Mangareva in pre-contact times.

and its country. The knotted cord also served Omai as a calendar, twenty-eight to thirty knots for each moon – four moons since they left Huaheine. (Clark 1969:20)

Ma'i received English lessons of a more formal nature during his stay in England, when his illiteracy and poor command of the language became of concern to a group of London philanthropists who felt that it would be irresponsible to let him return home without having taught him how to read and write (McCormick 1977:167). Their ulterior motive was evangelical. Through reading and writing, Ma'i could be introduced to the Christian faith and thus be entrusted with the spreading of the Gospel among his fellow countrymen on his return to Tahiti.

Granville Sharp obtained permission from the Admiralty to start the instruction. Ma'i attended a few classes and Sharp claimed that he had taught him the use of English letters and had made him sound every combination of vowels and consonants that letters are capable of. If we were to believe him, Ma'i was even able to write a letter to Dr Solander.

After 15 lessons Ma'i must have been bored to death. He pretended to have too many social engagements to attend further classes. It was probably the set textbook and Sharp's approach to 'English as a second language' that had turned him off. The phonetic section, written by Sharp himself and adapted to the needs of Ma'i, is entitled 'An English Alphabet, for the use of Foreigners: wherein the Pronunciation of the Vowels, or Voice-Letters, is explained in Twelve short general Rules, With their several Exceptions, As Abridged (For the Instruction of Omai) From a larger Work'. It contains rules of didactic transparency like the following: "The vowels are pronounced short in all syllables ending with a consonant, (except in the particular cases hereafter noted,) and the three first vowels have the foreign articulation, without any other material difference except that of being pronounced short" (McCormick 1977:167).

## 6. IS 'OTAHEITAN' AN EASY LANGUAGE?

While philosophers and literati looked at the Tahitian language and the linguistic ability of its speakers from a sociophilosophical point of view in an attempt to find an answer to the rather elusive question whether man in a state of nature was superior to his civilised cousin, there were other people who had a very pragmatic interest in Tahitian. These were the missionaries of the London Missionary Society who were planning to save the Noble Savages from eternal damnation by converting them to the Christian faith. From what they could read in the published works and from what they concluded observing the limited linguistic skills of the South Sea Islanders, it would be very easy to learn Tahitian for somebody who spoke such a sophisticated tongue as English. Had not a simple corporal of the marines on Cook's first voyage been able to speak the language fluently after only a three months' stay in Tahiti? And had not a trustworthy person with first hand experience like Captain Bligh assured Reverend Haweis, who was organising the missionary enterprise in the South Seas that the learning of a Polynesian language would present no difficulty? (Newbury, ed. 1961:XXIX).

It seems almost an act of higher justice that the roles were reversed in 1798 when the first English missionaries arrived in Tahiti and tried to learn the language. Their difficulties were such that for years the whole project looked like it would end in a disastrous failure. On 10 April 1799, Jefferson wrote in his journal:

Our growth in knowledge of the language is still slow and in many cases uncertain; which is in great measure owing to our not being able to catch the sound of the words, with that exactness that is necessary; there is as great a labour to arrive at the true sense and

meaning of a word, or its various meaning; for one word is used to express very opposite things in different sentences. (Newbury, ed. 1961:XLIV)

And years later, in 1806, the missionaries recognised that they had been gulled into believing that Tahitian was an easy language to learn.

It has been represented as uncommonly easy of attainment; but we know the contrary by long experience. The Tahitian language, as may justly be expected, is destitute of all such words, common among civilized nations as relate to the arts and sciences, law proceedings, trade and commerce, and most of those made use of in Theology &c. But when viewed as having relation to objects known, and in use among the natives it is full and copious. (Newbury, ed. 1961:85)

By then they also knew that they had made fools of themselves in the early days when they used to say to the Tahitian king, "*Mity po, tuaana*". By putting together *mity* [maita'i] 'good', *po* [po:] 'night' and *tuaana* [tua'ana] 'elder brother' they had produced an interlanguage construct which easily matched in its naiveté any Omaism<sup>8</sup>. When the Tahitians finally found out what the missionaries wanted to express, they just laughed without making any assumptions about speech defects or inferior mental capacity of the religious envoys. After all, nobody in paradise had ever heard of Rousseau.

## BIBLIOGRAPHY

- ALEXANDER, M., 1977, *Omai 'Noble Savage'*. London: Collins & Harvill Press.
- BEAGLEHOLE, J.C., ed., 1959, *The Endeavour journal of Joseph Banks 1768-1771*. 2 vols. Sydney: Trustees of the Public Library of NSW, Angus and Robertson.
- BOUGAINVILLE, L.- A. de, 1772, *A voyage around the world*, translated from the French by J.R. Forster. London: Nourse.
- CHINARD, G., ed., 1935, *Didérot. Supplément au Voyage de Bougainville*. Paris: Librairie E. Droz.
- CLARK, T.B., 1969, *Omai, the first Polynesian ambassador to England*. Honolulu: University of Hawaii Press.
- ELLIS, W., 1969, *Polynesian researches: Society Islands*. Rutland, Vermont, and Tokyo: Charles E. Tuttle Co. [Originally published in 1829.]
- FAIRCHILD, H.N., 1961, *The noble savage: a study in romantic naturalism*. New York: Russell & Russell.
- KNOWLTON, E., 1955, A letter from Tahiti, 1769. *Journal of Oriental Literature*, 8-15.
- MAZELLIER, P., ed., 1978, *Le mémorial polynésien*. Papeete: Hibiscus Editions.
- McCORMICK, E.H., 1977, *Omai: Pacific envoy*. Auckland: Auckland University Press.
- MONTAIGNE, M., 1894, Of the canibales. In Henry Morley, ed. *The essays of Michael Lord of Montaigne*, 92-99. Translated by John Florio. London: George Routledge.
- NEWBURY, C.W., ed., 1961, *The history of the Tahitian Mission 1799-1830*. Cambridge: CUP.
- PARKINSON, S., 1773, *A journal of a voyage to the South Seas...* London. (1972 facsimile, Adelaide.)
- PEACOCK, D.J., 1987, *Lee Boo of Belau: a prince in London*. Honolulu: University of Hawaii Press.
- ROUSSEAU, J.-J., 1974, *Essai sur l'origine des langues*. Paris: Aubier Montaigne. [First published 1754.]
- TAILLEMITE, E., 1977, *Bougainville et ses compagnons autour du monde: 1766-1769, journaux de navigation*. 2 vols. Paris: Imprimerie Nationale.
- TINKER, C.B., 1964, *Nature's simple plan*. New York: Gordon Press.

<sup>8</sup>If they understood *mity* as their word *maitai*, the phrase would be an assertion to this effect, 'Good (is the) night, brother'. The simple declaration, 'Good-night, brother', would be unintelligible to the Tahitians, though the language were correct; a corresponding wish among them would be, *la ora na oe i teie nei po* (Ellis 1969:14).



LOSS OF FINAL CONSONANTS  
IN THE NORTH OF NEW CALEDONIA

J.-C. RIVIERRE

1. INTRODUCTION

M. Leenhardt (1946:xx) divides the languages of mainland New Caledonia into two main groups: the Southern group and the Northern group, "separated [in the centre of the mainland] by a line which can be drawn between Ponérihouen and Koné". This line therefore separates Paicî (PAC), classed among the Northern languages, from Ajië (AJE), considered as a Southern group language (see the map and list of abbreviations therein).

Leenhardt justifies the positioning of this boundary by enumerating a number of cultural features typical of each of the areas, but on the linguistic level only gives a few indications, which are not very convincing. Nevertheless, this classification is not without interest since, as Haudricourt (1965) has shown, all the languages situated to the north of this line, including Paicî, have undergone deep and similar transformations in their consonant orders, whereas, as far as this feature is concerned, the Southern languages have remained more stable. Most Northern languages have developed consonant systems rich in aspirates, fricatives and even postnasals, whereas the Southern languages have instead developed a rich vowel system, which contains an average of thirty or so units.

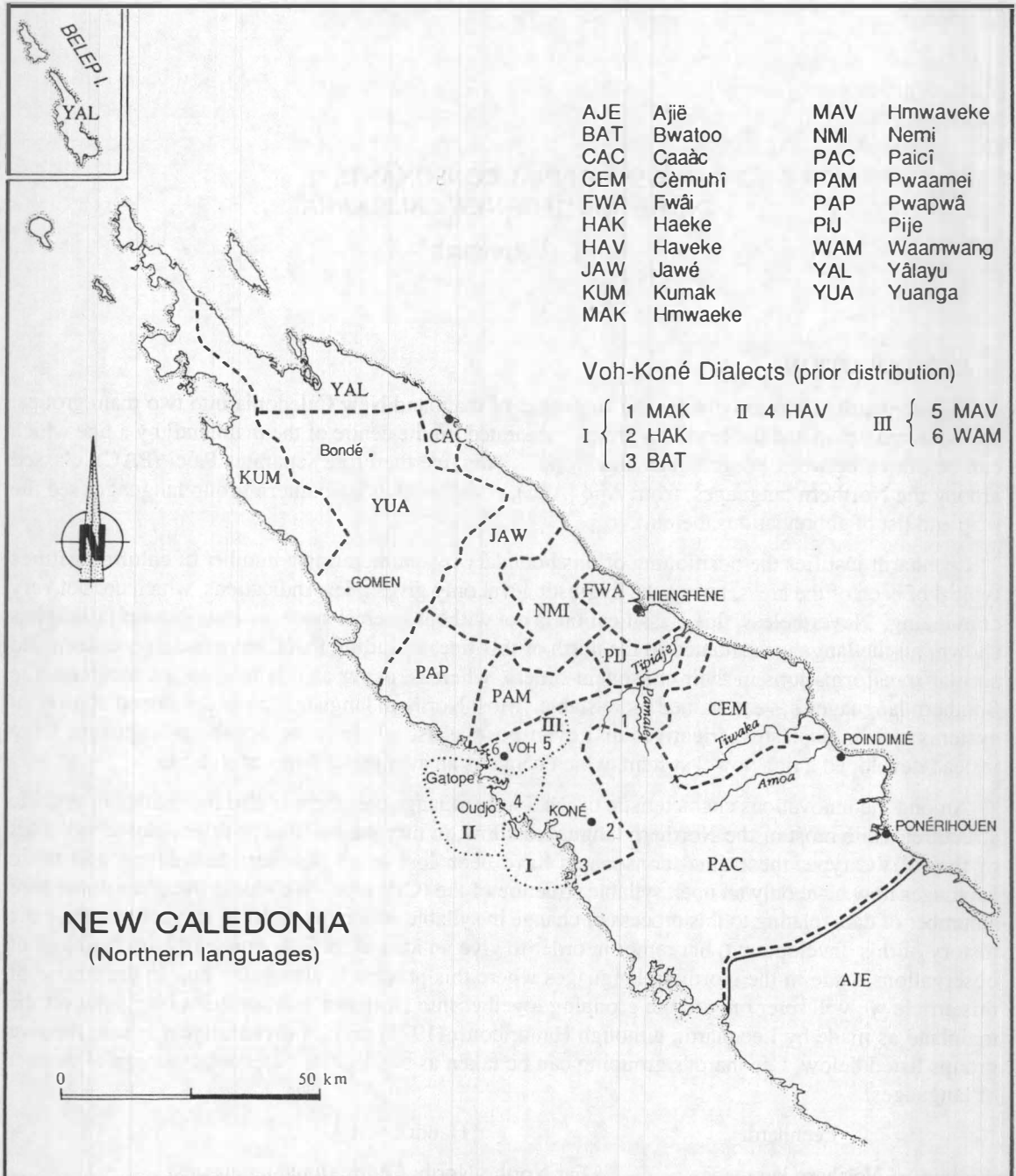
Among the innovations characteristic of the Southern languages there is also the change in syllabic structure: while most of the Northern languages allow, in morpheme-final position, closed syllables of the (C)VC type, these final consonants have been lost in all Southern languages, and these languages now have only an open syllabic structure of the (C)V type. We would like to examine here a number of data relating to this process of change in syllabic structure, not in an attempt to retrace the history of this development, but rather in order to give an idea of its full complexity, on the basis of observations made in the Northern languages where this process is also going on. In the course of this article we will refer only to the grouping together into Northern and Southern languages on the mainland as made by Leenhardt, although Haudricourt (1971) divides these languages into the five groups listed below. Leenhardt's grouping can be taken as the highest order subgrouping of this set of languages.

Leenhardt	Haudricourt
Northern languages	Far North ; North; Centre (tonal languages)
Southern languages	South ; Far South (tonal languages)

Robert Blust, ed. *Currents in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, 415-432. *Pacific Linguistics*, C-117, 1991.

© J.-C. Rivierre





## 1.1 PLAN

In this article we would like to emphasise that the loss of these final consonants (henceforth FCs), which depends partly on internal phonological factors, is also largely determined by the effects of contact and of spreading from linguistic areas in which these developments have already taken place. The southernmost languages and dialects of the Northern area (Paicî, Cèmuhi, the Voh-Koné dialects), spoken in the neighbourhood of Ajië and of the other Southern languages give a good illustration of this:

### (a) PAC (Paicî)

We shall see that the loss of FCs has been fully achieved in PAC, bringing about vocalic innovations similar to those that can be observed in all Southern languages. In its syllabic structure as well as in its vowel system, PAC has become similar to the Southern languages immediately adjoining it.

### (b) CEM (Cèmuhi)

PAC in its turn has influenced CEM, spoken on the north-eastern side of the mainland, bringing about the loss of certain FCs. CEM, with its incomplete system of FCs, represents an intermediate stage between the Hienghène languages, which have remained conservative, and PAC, which has become an open-syllable language.

### (c) KON (the Voh-Koné dialects)

On the north-western side, the dialects from the Voh-Koné region show an identical progression. We can distinguish here between a front of 'developed' dialects spoken in contact with open-syllable languages, and, further back from this zone, a set of more conservative dialects which are gradually being affected by this development.

## 2. LOSS OF FINAL CONSONANTS IN PAC

PAC and AJE differ considerably as far as their lexis, grammar and even prosody is concerned, since PAC has developed a complex morpho-tonal system, whereas AJE has not. Moreover, as we have pointed out, these two languages have undergone a divergent evolution in their initial consonant system:

- Whereas in AJE and in the Southern languages only the point of articulation of former uvulars has been advanced, this movement has been extended to velars and palatals in all the Northern languages. Former velars and uvulars, affected by this advance, became weakened and subsequently disappeared in PAC as well as in CEM, the two tonal languages from the Northern area.
- Most of the Northern languages, unlike their Southern counterparts, have also been affected by palatalisation of their former dental consonants.

These changes are summed up in Table 1 (after Haudricourt 1965).

TABLE 1: REFLEXES OF POC INITIAL CONSONANTS IN NEW CALEDONIA

Northern area (including PAC)		AJE(South)
* uvular	>	velar (> Ø)
* velar	>	palatal (> Ø)
* palatal	>	apical
* retroflex	>	apical
* dental	>	apical / palatal / velar
* bilabial	>	bilabial / velar

This differentiation between PAC and AJE, languages situated on either side of the borderline between the two main language groups, has however been mitigated by a number of evolutionary similarities:

- Identical treatment of certain consonant orders, in particular the old geminate velar and uvular consonants now represented by Ø in both languages (\**qq* > Ø, \**kk* > Ø, cf. examples (1), (3) and (8) in Table 2).
- Above all, like AJE and the Southern languages, PAC has become an open-syllable language, possessing a diversified vowel system, rich in nasal vowels. For the most part, these syllabic and vocalic similarities have the same origin: FCs, which are still present in the North, have been lost in all Southern languages and in PAC, bringing about an enriching of the vowel system, since the loss of oral consonants was followed by the appearance of new vowel timbres, while the loss of nasal consonants played an obvious part in the genesis of nasal vowels, as we can see from examples (8), (9) and (10).

Other developments can explain the appearance of nasal vowels in PAC: some of these are shared with the Northern languages, in particular, nasalisations resulting from the simplification of postnasalised consonants (*C<sup>n</sup>V* > *C<sup>~</sup>V*); others, however, are shared with the Southern languages, such as bilateral nasalisations arising from *-n-* in intervocalic position (*V<sup>n</sup>V* > *ṼrṼ*). Cf. example (6).

TABLE 2: DIVERGENT EVOLUTION OF NORTHERN AND SOUTHERN LANGUAGES

	POC	KUM	NORTH				SOUTH
			NMI	KON (MAV)	CEM	PAC	AJE
(1) <i>Montrouzierra sp.</i>		<i>hup</i>	<i>hup</i>	<i>xup</i>	<i>úp</i>	<i>ú</i>	<i>u</i>
(2) ashes	* <i>ndapu</i>	<i>daap</i>	<i>dap</i>	<i>dap</i>	<i>dēp</i>	<i>dāl</i>	<i>de</i>
(3) pole for yam vines			<i>havit</i>	<i>xait</i>	<i>áit</i>	<i>ápi</i>	<i>ai</i>
(4) flood		<i>kaavak</i>	<i>davec</i>	<i>javac</i>	<i>jāa</i>	<i>jēpe</i>	<i>dee</i>
(5) to die	* <i>mate</i>	<i>maak</i>	<i>mac</i>	<i>mec</i>	<i>mēt</i>	<i>bāl</i>	<i>me</i>
(6) bird	* <i>manuk</i>	<i>mālic</i>	<i>manik</i>	<i>manik</i>	<i>mēnī</i>	<i>mālṛl</i>	<i>mūrūr</i>
(7) year	* <i>taqu</i>	<i>ka</i>	<i>tok</i>	<i>jok</i>	<i>jō</i>	<i>jā</i>	( <i>ne-</i> ) <i>do</i>
(8) cultivate	* <i>quma</i>		<i>hum</i>	<i>xum</i>	<i>úm</i>	<i>ũ</i>	<i>ũ</i>
(9) turtle	* <i>popu</i>	<i>wan</i>	<i>pwen</i>	<i>vwen</i>	<i>pwēn</i>	<i>pwāl</i>	<i>pwē</i>
(10) breadfruit	* <i>kuluR</i>	<i>cen</i>	<i>cin</i>	<i>ḍiṅ</i> (BAT)	<i>īn</i>	<i>ĩ</i>	<i>kāl</i>

NB - In the lists of examples given in this paper, prenasalised consonants have been transcribed as voiced oral occlusives.

To resume, if PAC is definitely to be classed as a Northern language by virtue of the development of its initial consonants as well as in many other features, it also presents similarities with AJE and the Southern languages in its syllabic and vocalic structure. It would seem reasonable to ascribe these similarities to contact and to the spreading of the main evolutionary feature of the Southern languages: loss of FCs. Although this spreading and the traces left by this evolution now appear clearly to us, it is difficult, looking back, to reconstruct exactly how and in what order these changes came about. However, the current spreading of this feature towards the Northern languages can give us some clues as to how it took place.

### 3. LOSS OF FINAL CONSONANTS IN CEM

#### 3.1 FINAL CONSONANTS IN THE NORTHERN LANGUAGES

The system of FCs in the Northern languages is reduced to two sets, oral and nasal, divided into four orders of articulation:

-m	-n	-ɲ	-ŋ
-p	-t	-c	-k

These FCs correspond to either of the following:<sup>1</sup>

(a) former POC final consonants :

* <i>tumpuq</i> 'to grow; swollen'	PIJ, FWA, NMI, PAM PAP	<i>higuk</i> <i>chiguk</i>
------------------------------------	---------------------------	-------------------------------

(b) former intervocalic consonants which have become final:

* <i>taqu</i> 'year'	PIJ, FWA, NMI2 NMI1 JAW MAV	<i>tok</i> <i>dak</i> <i>jak</i> <i>jok</i>
----------------------	--------------------------------------	--

Ozanne-Rivierre (1982:55) notes that this trend towards closed syllables mainly affects morphemes without enclitics: "Independent nouns (without possessive suffixes), intransitive verbs and grammatical suffixes tend to end in a consonant (loss of the former final vowel, persistence of the former final consonant)":

-CV	→	-C
-CVC	→	-CVC

In addition to these etymological FCs there are others, non-etymological, which have been studied by Grace (1972), by Ozanne-Rivierre (1982) and, in the present volume, by Ozanne-Rivierre and Hollyman.

<sup>1</sup> As far as final oral consonants are concerned, former uvulars and velars have generally merged into -k (or Ø), except in CAC, KUM and PAP, where velars have become palatals (-c). The old dentals are now palatals (-c), except in KUM where they have become velars (-k).

### 3.2 FINAL CONSONANTS IN CEM

CEM, which is spoken at the southernmost limit of the languages which have preserved FCs, represents a transitional stage between these conservative languages and open-syllable PAC. CEM, much more so than PAC, is germane to the Northern languages by virtue of certain phonological features (a simple vowel system, FCs), and by the high percentage of common words shared with both the Voh-Koné dialects from the west coast and the languages spoken in the Hienghène area.

CEM, however, is also the closest language to PAC both in linguistic terms and on a social and geographical level:

- The two languages have an almost identical consonant system; they share the feature of having developed tones and are also very close grammatically.
- A great many PAC and CEM speakers are in contact in the vicinity of the Amoa valley, and the bilingual zone spreads from the Poindimié region (PAC area) as far as the Tiwaka (CEM area).

CEM, as an interlinking language, is of special interest, since it undoubtedly owes to its geographical position and to its contact with PAC the loss of several of its FCs. Its present system, incomplete in several places, shows that the process of dismantling of FCs is going on, and we should be able to find indications of this in its phonology.

The FCs system of CEM is currently as follows:

-m	-n	-	-ŋ
-p	-t	-	-

This table shows both the greater vulnerability of back FCs (velars and palatals) and the greater resilience of nasal FCs, since the nasal velar [ŋ] has been retained.

Many facts observable at the present time lend support to the idea that FCs in CEM are becoming weakened.

#### 3.2.1 UNSTRESSED FINAL SYLLABLES

In a polysyllable of the CVCVC type, the last syllable can be equated with an unstressed syllable. In nine words out of ten this syllable has no tone of its own; it takes on the same tone as the initial syllable when the word is pronounced in context. This final syllable is also liable to neutralisations:

- in consonants: the FCs system is more limited than that of initial or intervocalic consonants.
- in vowels: this syllable always contains fewer distinctions of aperture than the first (cf. Rivierre 1980:44).

#### 3.2.2 VARIATION IN FINAL CONSONANTS

##### 3.2.2.1 CLOSED/OPENED SYLLABLES ALTERNATIONS

In a large number of contexts, the final oral consonants *-p* and *-t* alternate with the syllabic forms *-V* or *-CV*. These alternations occur in front of an affix or when the following word begins with a consonant.



These alternations are free when the word is pronounced in isolation.

-p alternates	with -ε	āpulip ~ āpulie	man
	with -o	if the preceding vowel is a back vowel: būp ~ būo	to burst
-t alternates	with -le	būput ~ būpule	bracelet

### 3.2.2.2 ORAL/NASAL ALTERNATIONS

Oral and nasal FCs with the same point of articulation alternate in a number of words, as if the opposition between the two features were tending to become blurred:

- p alternates with -m
- t alternates with -n

These variations are attested in the bilingual zone situated between the Tiwaka and Amoa valleys, close to the PAC area. The morphemes affected by this feature generally have a final oral consonant in the conservative zone, situated to the north of the Tiwaka. To these final oral consonants also correspond nasal finals which do not alternate further south. We find, for example:

	northern CEM	southern CEM
ash	dēp	dēp ~ dēm
to have fun	lúup	lúup ~ luum
beam	pwāhēdop	pwāhēdom
fat	jēp	jēm
missionary (Fr. père)	pēt	pèn

These correspondences therefore affect older words (\*ndapu 'ash') as well as more recent borrowings (Fr. père 'father' → pēt, pèn). In both cases we see that a nasal consonant can replace an oral one in morpheme-final position. We may also notice that these alternations are more frequent among bilabial consonants than among apicals.

## 3.3 CAUSES OF THE LOSS OF FINAL CONSONANTS

All these facts taken together enable us to formulate certain hypotheses as to the causes of this change in the structure, and how it is carried out.

### 3.3.1 EXTERNAL CAUSES

The examination of FCs in CEM confirms the part played by contact in the development and dislocation of the final consonant system. This influence is revealed in the gradual transition from a 'developed' zone with open syllables, in contact with the Southern languages (PAC), through an intermediate zone in which we find a reduction in the number of FCs (CEM), and finally to the conservative zone of the Northern languages.

This progression on a regional scale is echoed on an individual scale within CEM itself. Going from south to north, we can distinguish:



- a borderline zone with PAC, near the Amoa, in which sporadic loss of FCs, bilabials or apicals, is noted;
- a transitional zone, characterised by alternations between oral and nasal consonants (*dɛp ~ dɛm*);
- the northern CEM area, the most conservative, situated outside the bilingual zone.

### 3.3.2 INTERNAL CAUSES

In addition to these external causes there are other internal factors, within the system itself, that contribute to the weakening of FCs:

- a tendency to avoid strings of consonants, as is evidenced by the alternation of final oral consonants and by the readjusting of the CVC-type syllabic sequence into CVV or CVCV (*būp ~ būɔ*, *bwēt ~ bwēle*).
- in polysyllables, FCs appear in unstressed position. More generally, initials prevail over finals, at the level of the syllable (CVC) as well as that of the morpheme (CVCVC).

### 3.3.3 THE PROCESS OF CHANGE

The following suppositions may be made concerning the process of loss of FCs:

- Back consonants – palatals and velars – are affected before front ones, bilabials and apicals.
- Oral consonants are more vulnerable than nasals. Several facts lend support to this hypothesis:
  - (a) The preservation of the nasal velar [ŋ] in the present phonological system;
  - (b) The alternation of the final oral consonants with -V or -CV. Forms with a final vowel might eventually replace forms with a final consonant : *būp ~ būɔ* → *būɔ*;
  - (c) Alternations and substitutions of FCs tend to increase the number of nasal finals : *dēp ~ dēm* 'ash' → *dēm* (cf. section 3.2.2.2).

These variations in FCs appear to indicate that they are becoming weakened, and thus foreshadow their disappearance. According to this hypothesis, the loss of a final consonant would not simultaneously affect all morphemes in which the consonant occurs. Starting off from a small portion of these lexemes, the process of loss would then spread gradually to the whole of the vocabulary concerned. This last supposition is backed up only to a slight extent by observable facts in CEM. However, if we were to find confirmation of this in the region currently under study, it would give further weight to the facts given by Chen and Wang (1975) in their explanation of the model of change by lexical spreading. This type of spreading can be observed in the dialects spoken in the Voh-Koné area.

#### 4. LOSS OF FINAL CONSONANTS IN THE VOH-KONÉ DIALECTS

##### 4.1 THE VOH-KONÉ DIALECTS

The Voh-Koné dialects represent an intermediate stage, on the west coast, between the conservative Northern languages and the open-syllable languages (PAC, AJE).

These dialects currently number four (MAV, MAK, HAV, HAK), since WAM, the Voh dialect which was still spoken in Leenhardt's time, has now disappeared. The disappearance of WAM was preceded historically by other recessions and upheavals in these dialect areas:

- PAC speakers gradually replaced HAK speakers in the footlands of the Koné region two or three centuries ago (cf. Bensa and Rivierre 1988);
- HAK and BAT have withdrawn from the coastal region to the south of Koné where, according to Leenhardt (1946:xv), they used to be in contact with AJE speakers. "As recently as 1910 all the old people of Koné could still understand AJE".
- MAK speakers, formerly settled in the central mountain chain (High Tipidjé, Pamalé) and on the western slopes, have been driven down to the nearest lowland valleys, on the west coast or on the east coast.<sup>2</sup>

Since these recessions, which were the result of pre-colonial wars and colonisation, the speakers of these four dialects have lived on a narrow stretch of the coast between Koné and Voh, as well as in the lowland valleys adjoining these two localities. They live there grouped together, and often mixed, in a small number of reservations, in which they nevertheless intend to preserve their dialectal differences to show their origin, their membership of a particular clan, and even their place in society.

The current distribution of these dialects therefore does not allow us to understand the influences they have been subjected to in the course of their history. We shall divide them into three groups here, on the basis of their traditional implantation and their degree of proximity to the 'developed' languages to the South and South-East (AJE, PAC, CEM; cf. map).

Group I comprises the dialects traditionally in contact with CEM or with open-syllable languages:

- MAK, formerly close to CEM and to PAC in the Voh hinterland;
- HAK and BAT, formerly in contact with PAC and AJE.

Group II: HAV, spoken on the sea coast between Voh and Koné, occupied an intermediate position, being the least directly exposed to outside influences, either from the North or from the South.

Group III comprises the WAM and MAV dialects spoken in the vicinity of the conservative Northern languages, PAM and PAP.

##### 4.2 THE FINAL CONSONANTS SYSTEM OF THE VOH-KONÉ DIALECTS

The FCs system characteristic of these dialects is identical to that which is to be found in other Northern languages. It consists of two sets, oral and nasal consonants, divided into four orders of articulation. Table 3 sums up, for each of these orders, the percentages of losses of FCs noted in the

<sup>2</sup> On the west coast, most MAK speakers are grouped together with MAV speakers in Tiéta village. The ones who have settled on the east coast now speak a slightly different dialect, known as Vamalé.

different dialects (WAM included, as Leenhardt 1946 gives a vocabulary list for this dialect). As we can see, these figures only concern the oral consonants, since final nasals have remained stable. On the whole, we find a situation which has by now become familiar: back consonants are lost before the others, and a change occurring in a given linguistic area tends to spread and to affect its closest neighbours.

In this case, the FCs change starts out from the Group I dialects (HAK, BAT and MAK) and reaches the Group III dialects (WAM, MAV), which are in contact with the conservative languages. These transitional dialects undergo gradual loss of FCs in their lexis, without, however, gaps appearing in their FCs system as a whole, as is the case in CEM. As we can see from Table 3, no FCs have completely disappeared, and each of the dialects from this region can still be considered as having a full set of FCs.

The process observed here is therefore less advanced than in CEM (see section 3). We are closer here to the origins of the process whose mechanism we are trying to imagine, and this allows us to gain a better insight into its progression and its spread.

TABLE 3: PERCENTAGES OF LOSSES OF FCs IN VOH-KONÉ DIALECTS

	III		II	I		
	WAM	MAV	HAV	HAK	BAT	MAK
-k (after u, o)	46 %	41 %	93 %	93 %	96 %	88 %
-k (after i, e, a)	8.5 %	11 %	40 %	83 %	66 %	80 %
-c	5 %	—	61.5 %	79 %	93 %	90 %
-t	—	—	—	—	—	—
-p	—	—	—	—	44 %	—

By presenting the percentages of losses of FCs in the different dialects in table form, we are able to confirm and to clarify some of the observations made earlier in CEM.

#### 4.3 SUCCESSIVE LOSSES OF DIFFERENT FCs

FCs appear to be lost in the following order:

- (a) Back consonants are the first to disappear, but *-k* seems to begin disappearing before *-c* does.

The percentages for the losses of these two consonants are more or less equal in Zone I, but in Zone III, where the WAM and MAV dialects are starting to lose their FCs, loss of *-k* appears much earlier than that of *-c*.

- (b) Final *-k*'s everywhere disappear in much greater number when they follow a back vowel than when they follow a front vowel.

We may then suppose that the process of loss of FCs is triggered by the disappearance of *-k* after [o], [u]. It does indeed seem likely that the final velar consonant is acoustically more vulnerable in this position.

(c) Judging by a number of indications in CEM *-p*'s disappear before *-t*'s.

This can be seen in BAT, where *-p* has become *-t* in almost half of the lexemes concerned. This evolution of *-p*, which is different from the evolution that can be seen in CEM, cannot be considered in this case as a commonplace event; on the other hand, the stability of *-t* compared to other consonants is in no way surprising.

The order in which the four FCs in the dialects begin to disappear would therefore seem to be as follows:

1	<i>-k</i>
2	<i>-c</i>
3	<i>-p</i>
4	<i>-t</i>

It is obvious from the figures given that these consonants do not disappear one after the other, with each one waiting for the previous one to have fully disappeared before beginning to disappear itself. The disappearance of each of these consonants is a process which takes place over a certain length of time, these periods of time being more or less concomitant or spaced out; the loss of *-k* seems rapidly to trigger that of *-c*, whereas (as is also the case in CEM) the loss of front consonants seems to take place at a later stage.

Similarly, there is no need for a final consonant to have fully disappeared for the innovation to begin to spread to neighbouring dialects. A long time before a consonant disappears, the pronunciation of certain lexemes affected by the innovation is adopted or imitated by speakers of other dialects, for complex reasons which would need to be researched in a separate linguistic and sociolinguistic survey.

It would nevertheless be simplistic to think that innovations spread and are transmitted unilaterally, from innovating languages to conservative ones. Although the loss of final *-k* and *-c* appears to be spreading to the whole of the Group I lexical stock in HAK, BAT and MAK, evidence shows that these dialects are continuing to borrow lexemes ending in these same consonants from conservative languages to the north. Some of these borrowings are easy to detect because initial *p-* and *k-*, which have remained stable in the conservative languages, have become respectively *v-* and *ɣ-* in the Voh-Koné dialects.

- We thus find *kolook* 'albino' in MAK; and the name for the corncrake, *pik*, also attested in this form in PAM, PAP and in the Hienghène languages.
- The name for the thrush (*Philemon diemenensis*) is *katholec* in BAT, from which almost all final *-c*'s have disappeared.
- Finally, although the change from *-p* to *-t* is still going on in BAT, borrowed lexemes with final *-p* are given either a final *-p* or *-t*, although the reasons governing these two different pronunciations are not clear:

BAT loan-words	with <i>-p</i>		with <i>-t</i>	
From Northern languages	<i>putep</i>	blackbird	<i>kapoot</i>	to burp
			<i>puuat</i>	to float
From European languages	<i>pwaip</i>	pipe		
	<i>taap</i>	table		

In spite of these repossessions through borrowing, the loss of these FCs seems gradually to be affecting all lexemes, and at the same time spreading beyond the dialect boundaries.

To conclude, we will look at two problems concerning:

- The nature of the change. Is the loss of an FC in a given lexeme gradual or sudden? Is the change adopted in an identical way in different uses of the same lexeme?
- How the change is adopted. Is the new pronunciation of a given lexeme (without the FCs) adopted simultaneously by all speakers, or does it gradually come into use in the community as a whole?

As one may imagine, an exhaustive study of the loss of FCs in any one dialect would ideally require plentiful materials (dictionaries, texts) giving for each lexeme transcriptions gathered in different contexts, from different informants and at different times. Of course, such materials are not available to us, but those that are available are nonetheless rich enough to provide us with a number of useful indications.

#### 4.4 VARIATIONS AND LOSS IN THE FINAL CONSONANT OF A LEXEME

Several different pronunciations for the same lexeme seem to coexist for a certain length of time and in different ways before the innovation is finally accepted.

Under item number 482 'flying fox' (*Pteropus*) in Leenhardt (1946), the following transcriptions are given:

	WAM (III)	MAV (III)	HAV (II)	HAK (I)	MAK (I)
flying fox ( <i>Pteropus</i> )	<i>methok</i> [ <i>meθok</i> ]	<i>mesok</i>	<i>metho</i> [ <i>meθo</i> ]	<i>metho</i> [ <i>meθo</i> ]	<i>metho</i>

These transcriptions show a discontinuity between the Zone III dialects, in which this item has a final *-k*, and the dialects from Zones I and II. How then are we to understand the transfer of this discontinuity to Zone III, that is, how does the possible transition from *mesok* to *meso* come about in MAV?

In fact, the list given above represents only a very approximate version of the real situation:

- Even in Zone II there still exists a residual pronunciation with final *-k*, attested for HAV in [*meθook*] 'angel fish', whose nickname is '*Pteropus*' in many mainland languages. The variety of referents to which a given term can apply can bring about variations in its signifier and to divergent evolutions, which may or may not be resolved afterwards.
- One need only refer to the other transcriptions given for this term in Zone III to see the following variations appear (from Leenhardt 1946:489):

	WAM (III)	MAV (III)	HAV (II)
482 flying fox ( <i>Pteropus</i> )	<i>meθok</i>	<i>mesok</i>	<i>meθo</i>
751 a male flying fox	<i>meθo xau</i>	<i>meso xayuk</i>	<i>meθo xau</i>
752 the female flying foxes	<i>i meθo thamo</i>	<i>ni meso thamo</i>	<i>ni meθo thamo</i>
753 some flying foxes	<i>be meθo</i>	<i>ni be meso</i>	<i>ni be meθo</i>



This example along with several others found in the Zone III dialects shows sporadic dropping of FCs before a consonant and a tendency to avoid strings of consonants; this practice, however, is not observed as a general rule in any of these dialects. Alongside these contextual variants, and no doubt in correlation with them, we find free variants with or without FCs when the word is pronounced before a pause (cf. MAV 482: *mesok*, 753: *be meso*).

In short, there is only an outward discontinuity between HAV (II) and the Zone III dialects, and this also applies to other items: a number of variations of this kind can be found in Leenhardt (1946), and they cannot all be put down to misprints or to inconsistencies in the transcriptions!

#### MAV (III)

conical meeting house	<i>poayu</i>	(139)	<i>poayuk</i>	(650)
two	<i>talo</i>	(768)	<i>thalok</i>	(888)
male	<i>xayu</i>	(104)	<i>xayuk</i>	(751)
sail	<i>ɲavo</i>	(413)	<i>ɲavok</i>	(418)

We may note that these variations seem to be frequent in the apparently vulnerable acoustic sector of final *-k*'s after a back vowel. These *-k*'s are elided in some cases before a consonant and are becoming less and less frequently sounded, since they have now disappeared from almost half the words in which they were previously attested in MAV (cf. Table 3).

By examining these variations, we are able to gain a better understanding of the way in which the complete loss of the consonant is likely to come about one day in the example given (*mesok*), as it is in all words which share this ending, under the pressure of both external factors (influence of Zone I and II dialects) and of internal systemic factors.

#### 4.5 SPREADING OF THE NEW PRONUNCIATION OF A LEXEME

An innovation is put forward by multilingual practices characteristic of a subgroup of speakers, and subsequently spreads to the dialectal community as a whole.

We are able to compare and contrast the results of several lexical surveys made for HAV (II) as it is spoken in Gatope and Oudjo. There is the Leenhardt survey (LN), which dates from 1939 and which was published in Leenhardt (1946), for dialect number 30 (Haveke), but which gives no indication as to the source of the data. The Haudricourt survey (HD) made in Oudjo in 1959 was a revision of the latter. Our own survey (RV), carried out in an improvised fashion in Oudjo in 1971 was not based on any document already existing. We should point out that it was made using one of the oldest informants in the area (a man who had fought in the First World War, born between 1890 and 1895).

The results of these three surveys agree to a great extent, and this confirms the durability of dialectal differences at the present day: a resilience which is surprising when one considers the grouping together of different speakers since the beginning of the century and the closeness of their linguistic systems.



TABLE 4: SOME MORPHEMES WITH FINAL *-k* IN HAV (II)

HAV			LN (1939)	HD (1959)	RV (1971)
bee	501	honey bee	<i>we mapuk</i>	—	<i>hmapu</i>
yam	302 303 304 633	yam top of yam bottom of yam first feast of the yam season	<i>wuvuk</i> <i>boan wuvuk</i> <i>hwatha wuvuk</i> <i>thoto wu</i> (= <i>wuvu, wuu</i> ?)	<i>yuwu</i>	<i>yuu</i>   <i>θoto yuu</i>
uselessly	1079 1144 1145	do it uselessly do it all the same in vain	<i>xaman xhayuk</i> <i>wa xhaiuk</i> <i>xaman xhau</i>	<i>yaman xayuk</i> <i>wa xauk</i> <i>yaman xau</i>	<i>xaau</i>
tree, wood	351 344 345 346 347	tree wood piece of wood twig pole	<i>ðek</i> <i>ðe</i> <i>bala ðe</i> <i>hmoauvan ðek</i> <i>du ðe</i>	<i>ðek</i> <i>ðek, ðe</i> <i>bala ðe</i>  <i>du ðe</i>	<i>ðee</i> <i>ðee</i> <i>baala ðee</i>  <i>duu ðee</i>

TABLE 5: SOME MORPHEMES WITH FINAL *-k* IN THE NORTHERN LANGUAGES

	NMI	PAP	PAM	WAM(III)	MAV(III)	LN	HAV(II) HD	RV	HAK(I)	BAT(I)	MAK(I)
wasp	<i>maguk</i>	<i>hmapuk</i>	<i>hmapuk</i>	<i>hmapu</i>	<i>hmapuk</i>	<i>hmapuk</i>	-	<i>hmapu</i>	<i>hmapu</i>	<i>hmapu</i>	<i>mapu</i>
yam	<i>kuuk</i>	<i>kuvic</i>	<i>kuvuk</i>	<i>wuvuk</i>	<i>uvuk</i>	<i>wuvuk</i> <i>wu</i>	<i>yuwu</i>	<i>yuu</i>	<i>yuvu</i>	<i>wuu</i>	<i>uvu</i>
uselessly	<i>xauk</i>	<i>xau</i>	<i>hayu</i>	<i>xauk</i>	<i>xayuk</i>	<i>xayuk</i> <i>xau</i>	<i>xayuk</i> <i>xau(k)</i>	<i>xaau</i>	<i>xau</i>	<i>xau</i>	<i>xayu</i>
tree	<i>ceek</i>	<i>ceec</i>	<i>yeek</i>	<i>ðeek</i>	<i>yeek</i>	<i>ðek</i> <i>ðe</i>	<i>ðek</i> <i>ðe</i>	<i>ðee</i>	<i>ðee</i>	-	<i>yee</i>

Several discrepancies do however appear in these three surveys when we examine the list of morphemes with final *-k* (cf. Table 4). We may note for example, that, as we have already seen in MAV (III), free variations appear in the transcriptions given respectively by Leenhardt and by Haudricourt. These notations seem to confirm the vulnerability of final *-k*'s, especially after a back vowel. While three instances of *-k* are still attested after [u] in Leenhardt's survey, these consonants in the same position all seem to have been lost thirty years later.<sup>3</sup>

As a matter of fact it may be erroneous to interpret the 1971 transcriptions as stemming from this evolutionary tendency. The comparison of these transcriptions with others made in neighbouring languages and dialects suggests another hypothesis (cf. Table 5): the differences noted are more or less contemporary individual variants, determined by the multilingualism of the speakers from the area. Alongside speakers who have family or political links with the conservative dialects of Zone III, there are others who are linked more closely to Zone I speakers. Developed pronunciations, without FCs, are introduced through these latter speakers, and they will coexist for a variable length of time alongside other uses, before finally becoming accepted by the community as a whole.

## 5. CONCLUSION

In the course of this article we have examined the process of loss of FCs as represented in the languages spoken in the southernmost part of the Northern half of mainland New Caledonia: those which are closest to the Southern half in which this loss has already been accomplished.

Alongside a number of internal factors which have contributed to this loss, the role played by contact appears to have been considerable since, on the whole, the closer the languages in question are to the centre of propagation, the more strongly they are affected by this process. The process of loss of any one consonant is in itself gradual, in two respects: the new pronunciation, without the FCs, seems to be initiated by a small number of speakers in a limited number of lexemes, then, step by step, the innovation gains on the community as a whole and gradually spreads to all the vocabulary concerned by the change.

But internal phonological factors conditioning the loss of FCs can also bring about this process by themselves, without the help of any external pressures. This is exactly what is happening in the Far North of the mainland, where the Yuanga (YUA) language, surrounded by languages with FCs, has become an open-syllable language, at least as far as the dialect spoken on the coastal area near Gomen (YUA-Gm) is concerned. In the more conservative variety of Yuanga spoken in the Bondé area (YUA-Bd), a final *-l* which corresponds to *-t* (< \**ns*, \**s*, \**d*) in the other languages is still attested among the final oral consonants. Opposite this single residual apical, the four nasal finals have been retained, and the FCs system for YUA-Bd is currently as follows:

FCs in YUA-Bd	<i>-m</i>	<i>-n</i>	<i>-ŋ</i>	<i>-ŋ</i>
(cf. Table 6)		<i>-l</i>		

The same process of loss of FCs has also begun in YAL, with the disappearance of the old back consonants \**-q* and \**-k*; but this development has come to a standstill, and, through borrowing back, YAL (which is spoken in the vicinity of YUA) currently possesses a full set of FCs.

<sup>3</sup>In fact the pronunciation *hmapuk* (with final *-k*) was recorded in the course of our survey (RV 1971) in *hmapuk-ca-we* 'jellyfish' (wasp-in-water).

TABLE 6: LOSS OF FINAL CONSONANTS IN YUANGA

		<i>*lauq</i> (PNC) fish	<i>*manuk</i> bird	<i>*tuqu(d)</i> to stand	<i>*qanso</i> sun	<i>*pukot</i> fishing net	<i>*ndapu</i> ashes
KUM		<i>nok</i>	<i>mālič</i>	<i>kuut</i>	<i>at</i>	<i>puiak</i>	<i>daap</i>
YUA	Gm	<i>nə</i>	<i>meni</i>	<i>kɔ</i>	<i>a</i>	<i>pwio</i>	<i>da</i>
	Bd	<i>nə</i>	<i>meni</i>	<i>kɔl</i>	<i>al</i>	<i>pwio</i>	<i>da</i>
YAL		<i>nō</i>	<i>māli</i>	<i>cur</i>	<i>gaar</i>	<i>pwiac</i>	<i>dap</i>
CAC		<i>nek</i>	<i>menic</i>	<i>coot</i>	<i>-gat</i>	<i>pwic</i>	<i>dep</i>
PAP		<i>nuk</i>	<i>manic</i>	<i>cuut</i>	<i>kat</i>	<i>pwiec</i>	<i>dap</i>
PAM		<i>nuk</i>	<i>manik</i>	<i>cuut</i>	<i>kat</i>	<i>puac</i>	<i>dap</i>
NMI		<i>nuk</i>	<i>manik</i>	<i>tuut</i>	<i>-gat/-kat</i>	<i>pwiec</i>	<i>dap</i>
KON (MAK)		<i>ɲu</i>	<i>mani</i>	<i>cuut</i>	<i>yat</i>	<i>vua</i>	<i>dap</i>

		<i>*wajka</i> boat	<i>Tyto</i> <i>alba</i>	<i>*kuluR</i> breadfruit	<i>*poɲu</i> turtle	<i>*lima</i> five	<i>*kakaumu</i> (PNC) to swim
KUM		<i>waaj</i>	<i>mwen</i>	<i>cen</i>	<i>wan</i>	<i>-nem</i>	<i>hyaa, hyao</i>
YUA	Gm	<i>wɔ</i>	<i>mwe</i>	<i>ci</i>	<i>pə</i>	<i>-ni</i>	<i>θɔ</i>
	Bd	<i>wɔɲ</i>	<i>mwen</i>	<i>cin</i>	<i>pən</i>	<i>-nim</i>	<i>ðɔm</i>
YAL		<i>waaj</i>	<i>mwen</i>	<i>yen</i>	<i>wan</i>	<i>nim</i>	<i>hyao</i>
CAC		<i>waj</i>	<i>mwen</i>	<i>cin</i>	<i>pwin</i>	<i>nim</i>	<i>jeum</i>
PAP		<i>waaj</i>			<i>pwen</i>	<i>nim</i>	<i>soom</i>
PAM		<i>waaj</i>	<i>mwen</i>		<i>pwen</i>	<i>nim</i>	<i>soom</i>
NMI		<i>waj</i>	<i>mwen</i>	<i>cin</i>	<i>pwen</i>	<i>nim</i>	<i>hyoom</i>
KON (MAK)		<i>waaj</i>	<i>mwen</i>	<i>in</i>	<i>vwen</i>	<i>nim</i>	<i>soom</i>

These developments, which are taking place independently from those that we have examined so far in this article, seem therefore to corroborate the latter, at least as far as the order of disappearance of FCs is concerned: they give further examples of the longevity of nasal and apical consonants, in contrast to the cases of early disappearance of back consonants attested in YAL. But these changes also show that the loss of FCs could, in a more or less distant future, spread and become generalised in the Northern group of languages.

We may then imagine the various hypotheses that future linguists, in the absence of any written documents, might put forward as to the loss of these FCs, attributing the innovation either to the proto-language of the Northern area or to the influence of the Southern area; seeing it as a late innovation generalised through spreading, or else as a parallel innovation at various points in the Northern area whereas, as we have seen, the real story of the disappearance involves, intricately linked together, several of these processes at the same time. There are then as many possibilities as we may legitimately conceive of, which justify the methodological precautions that Grace continually reminds us to take.

The problem turns out to be even more difficult when put on the scale of mainland New Caledonia as a whole since, as we have seen, some of the generalisations which can be made from the study of

the Northern group of languages do not apply to their Southern counterparts: the loss of FCs in the latter goes further back in time, and it took place in quite different phonetic and prosodic conditions, involving the transfer of features proper to FCs to the preceding segment or segments (cf. Ozanne-Rivierre and Rivierre 1989).

However, in contrast to the many problems posed by reconstruction studies, dialectology appears as a rich source of information, since it can sometimes enable us to retrace, within a given language area and in well-defined linguistic conditions, the successive stages, pauses and shifts in a linguistic change.

## BIBLIOGRAPHY

- BENSA, A. and J.-C. RIVIERRE**, 1988, De l'histoire des mythes. Narrations et polémiques autour du rocher Até (Nouvelle-Calédonie). *L'Homme* 106-107/XXVIII/2-3: 263-305.
- CHEN, M.Y. and W. S.-Y. WANG**, 1975, Sound change: actuation and implementation. *Language* 51: 255-281.
- CORNE, C.**, n.d., Dictionnaire yuanga-français. MS.
- GRACE, G.W.**, 1972, On canonical shapes in the language of Koumac. In Jacqueline M.C. Thomas and Lucien Berndt, eds *Langues et techniques, nature et société (Hommage à A.-G. Haudricourt)* 1: 135-140. Paris: Klincksieck.
- 1986, Hypotheses about the phonological history of the language of Canala, New Caledonia. In Chris Corne and Andrew Pawley, eds *Le coq et le cagou: essays on French & Pacific languages in honour of Jim Hollyman. Te Reo* 29: 55-67. Auckland: Linguistic Society of New Zealand.
- HAUDRICOURT, A.-G.**, 1963, *La langue des Nenemas et des Nigoumak (Dialectes de Poum et de Koumac, Nouvelle-Calédonie)*. Te Reo Monographs. Auckland: Linguistic Society of New Zealand.
- 1965, Problems of Austronesian comparative philology. In G.B. Milner and E.J.A. Henderson, eds *Indo-Pacific linguistic studies. Part I: Historical studies*, 315-329. Amsterdam: North-Holland.
- 1971, New Caledonia and the Loyalty Islands. In T.A. Sebeok, ed. *Current trends in linguistics*, vol. 8: *Linguistics in Oceania*, 359-396. The Hague: Mouton.
- n.d., Dictionnaires pwapwâ-français, pwaamei-français, yuanga-français. Lexiques des dialectes de Koné. MS.
- HAUDRICOURT, A.-G. and F. OZANNE-RIVIERRE**, 1982, *Dictionnaire thématique des langues de la région de Hienghène (New Caledonia): pije, fwâi, nemi, jawe*. Paris: SELAF.
- HOLLYMAN, K.J.**, n.d., Dictionnaire caaàc-français. MS.
- in this volume, Personalised and non-personalised possession: final consonants in Kumak and other languages of Far Northern New Caledonia.
- LA FONTINELLE, J. de**, n.d., Dictionnaire ajië-français. MS.
- LEENHARDT, M.**, 1946, *Langues et dialectes de l'Austro-Mélanésie*. Paris: Institut d'Ethnologie (Travaux et Mémoires XLVII).
- MICHAILOVSKY, B.**, n.d., Lexique yuanga-français. MS.
- OZANNE-RIVIERRE, F.**, 1982, Langues de Hienghène et proto-océanien: phonologie comparée. In Haudricourt and Ozanne-Rivierre, 1982:9-61.
- in this volume, Incorporation of genitive relators in the languages of New Caledonia and the Loyalty Islands.
- OZANNE-RIVIERRE, F. and M. MAZAUDON**, 1986, *Dictionnaire nyâlayu-français (Balade), Nouvelle-Calédonie*. Paris: LACITO-CNRS.
- OZANNE-RIVIERRE, F. and J.-C. RIVIERRE**, 1989, Nasalization/oralization: nasal vowel development and consonant shifts in New Caledonian languages. In Ray Harlow and Robin Hooper, eds *VICAL I: Oceanic linguistics: Papers from the Fifth International Conference on Austronesian Linguistics*, 413-432. Auckland: Linguistic Society of New Zealand.

**RIVIERRE, J.-C.**, 1972, Les tons de la langue de Touho (Nouvelle-Calédonie) : étude diachronique. *Bulletin de la Société de Linguistique de Paris* 67/1: 301-316.

1980, *La langue de Touho: phonologie et grammaire du cèmuhî (Nouvelle-Calédonie)* Paris: SELAF.

n.d., Lexique haveke-français. MS.

forthcoming, *Dictionnaire cèmuhî-français*. Agence pour le Développement de la Culture Kanak. Nouméa.

forthcoming, Contact-induced phonological complexification in New Caledonia.

HOW CONSERVATIVE ARE SEDENTARY LANGUAGES?  
EVIDENCE FROM WESTERN MELANESIA

M.D. ROSS

While I was working on the reconstruction of the prehistory of the Austronesian languages of western Melanesia (reported in Ross 1988), I noticed a curious distributional phenomenon. A large part of that work entailed the establishment of what Dyen (1956) calls 'minimal groups' of languages – 'minimal' in the sense that the languages within the group were more closely related to each other than to any languages outside the group. Quite often, I noticed that the languages within a minimal group were spaced out along a cline from the most conservative language to the least conservative and, moreover, that this cline often corresponded with the geographical distribution of the languages: if, for example, the languages of the group were strung out along a coastline, the most conservative was likely to be near one end of the string, and the least conservative near the other end. By 'most conservative' here, I mean 'having undergone the fewest phonological and morphosyntactic innovations relative to the proto-language ancestral to the minimal group'. Besides this correlation between innovativeness and geography, certain other correlations also emerged, and I became increasingly convinced that the most conservative language in a minimal group was often (but not always) the most sedentary, that is, one which had remained at its present location for the longest period of time, and, conversely, that the most innovative was often the one which had moved furthest from the 'home' location of the group.<sup>1</sup> In this paper I want to examine this conviction and its potential application in the process of reconstructing linguistic prehistory.<sup>2</sup>

It is probably easier to conceptualise the ramifications of the scenario I have just suggested if we work with a real example. The Schouten chain is a minimal group of languages whose speakers live

---

<sup>1</sup>I became convinced enough of the correlation between conservatism and sedentariness to hint at it in Ross (1988: 188). I am grateful to Frank Lichtenberk for drawing my attention to the fact that the assumption of this correlation needed to be made explicit and to be justified.

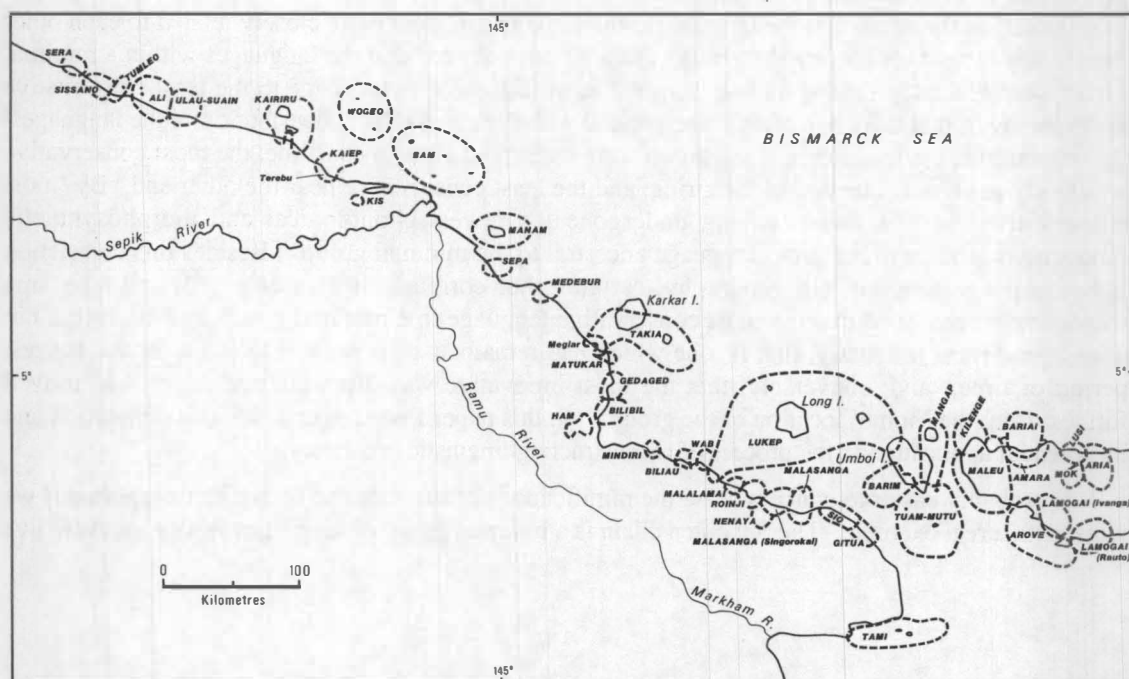
<sup>2</sup>It is a pleasure to offer this paper in honour of George Grace, who, as the references in the paper indicate, has been a formative influence on my thinking about Oceanic linguistic prehistory, as well as a source of encouragement in my work. I am grateful to Vladimir Belikov, Harold Koch and Matthew Spriggs for their helpful comments on an earlier version of this paper.



on the offshore Schouten islands and in coastal enclaves along the western half of the north coast of Papua New Guinea (see map). The languages of the chain are (from east to west):<sup>3</sup>

Medebur	Kairiru
Manam and Sepa (related at dialect level)	Ulau-Suain
Bam	Ali
Kis	Tumleo
Wogeo	Sissano
Kaiep and Terebu (related at dialect level)	Sera

The most conservative language of the group is Manam, the most innovative Sera, and the geographical positions of the languages between them more or less match their positions on a cline from the conservatism of Manam to the innovativeness of Sera. The one language which stands outside the pattern is Medebur.



AUSTRONESIAN LANGUAGES OF THE NORTH COAST OF PAPUA NEW GUINEA

<sup>3</sup>For more detailed information about the languages of the Schouten chain and for other references to them, see Ross (1988:121-132). Since Sepa and Terebu are related at dialect level to Manam and Kaiep respectively, they will not be considered further in this paper.

Since the Schouten chain is a minimal group, we know that all the languages in it are descended from one earlier language, Proto Schouten. The languages of the Schouten chain also belong to certain superordinate groupings, the best established of which is the very large minimal group of 'Oceanic' languages, which includes most of the Austronesian languages of the Pacific Ocean. All Oceanic languages, including those of the Schouten chain (and, at an earlier time, Proto Schouten) are descended from Proto Oceanic.

The fact that all the languages of the Schouten chain are descended from Proto Schouten is manifested in the fact that all of them share a set of innovations in common relative to Proto Oceanic (Ross 1988:124-125, 129-130). These are the innovations which Proto Schouten (or a forerunner) had undergone by the time of its break-up. The innovations which are of greater interest here, however, are those which the individual languages of the chain have undergone *since* the break-up of Proto Schouten, as these innovations are the measure of the conservatism/innovativeness of each language.

The phonological and morphosyntactic innovations which have occurred relative to Proto Schouten in each of the Schouten languages are shown in Tables 1 and 2. Note that, although Proto Schouten is a reconstructed language, it is quite a secure reconstruction, as it is based not only on the evidence of the Schouten languages themselves, but also on higher-order reconstructions (including Proto Oceanic) derived from data from numerous other Oceanic (and non-Oceanic Austronesian) languages. Hence there is nothing circular about the argumentation on which the tables are based.

TABLE 1: PHONOLOGICAL INNOVATIONS IN SCHOUTEN LANGUAGES

	1	2	3	4	5	6	7	8	9	10i	10m
Medebur	-	-	-	-	-	a	a	+	-	a	a, b
Manam	-	-	-	-	-	b	-	+	-	a	a, b
Bam	-	-	-	-	-	b	-	-	-	a	a, b
Kis	-	-	-	(+)	-	(b)	-	+	-	a, b, c	a, b
Wogeo	-	-	a	-	-	(b)	-	-	-	a	a, b
Kaiep	-	-	(a)	-	-	c	(b)	-	+	a, c	.
Kairiru	-	-	b	(+)	-	c	-	-	+	a, c'	a
Ulau-Suain	+	-	-	-	-	d	-	-	-	a	a, c
Ali	+	+	a'	-	a	d	c	-	+	a	a, c'
Tumleo	+	+	a'	+	(a')	d	c	-	+	a	a
Sissano	+	+	a'	+	(a')	d	c'	+	+	a	a
Sera	+	+	a'	+	a'	d	c'	-	+	a, b	a

KEY:

1, 2 etc. represent starting points for innovations; the numbering does not imply chronological sequence; 10i and 10m are word-initial and -medial environments of a single proto-phoneme.

a and b indicate alternate members of a set of innovations (see below).

a' indicates a sequence of two innovations: first a, then a further innovation applying to the output of a.

+ The innovation has occurred in this language.

- The innovation has not occurred in this language.

.

The data are insufficient to determine whether the innovation has occurred in this language.

(x) Innovation x is incomplete in this language.

x, y x and y have occurred in different lexical items.

\* marks Proto Schouten form.

1. \*v- > Ø-/\_a
2. \*b > p
3. a. \*-d- > -r-  
a'. \*d > r  
b. \*d > f̄
4. \*r > l (merger with \*l)
5. a. \*s > h  
a'. \*s > h > Ø
6. a. \*z > j, y (merger with \*j)  
b. \*z > r (merger with \*r)  
c. \*z > y, Ø  
d. \*z > s
7. a. \*j > j, y (merger with \*j)  
b. \*j > s  
c. \*j > c  
c'. \*j > c > t
8. \*ñ > n (merger with \*n)
9. \*g > k
- 10i. a. \*q- > Ø-  
b. \*q- > k-  
c. \*q- > g-  
c'. \*q- > g- > k-
- 10m. a. \*-q- > -Ø-  
b. \*-q- > -k-  
c. \*-q- > -g-  
c'. \*-q- > -g- > -k-

TABLE 2: MORPHOSYNTACTIC INNOVATIONS IN SCHOUTEN LANGUAGES

	11	12	13	14	15	16	17	18	19	20	21
Medebur	-	-	-	-	-	-	a	+	-	+	a'
Manam	-	-	-	-	-	-	-	-	-	-	-
Bam	-	-	a	-	-	-	a'	.	-	+	.
Kis	-	-	...	-	a'''	-	-	+	-	+	a'
Wogeo	-	-	a	a	-	-	a'	-	-	+	-
Kaiep	-	-	.	.	-	.	-	.	-	+	.
Kairiru	(+)	-	(a)	a	-	a	-	+	-	+	a
Ulau-Suain	+	+	a'	a'	a''	a'	a'	+	-	+	a'
Ali	+	+	a'	a'	a'	a''	-	+	(+)	+	a'
Tumleo	+	+	a'	a'	a	a''	-	+	(+)	+	a'
Sissano	+	+	a'	a'	a''	a''	a''	+	+	+	a'
Sera	+	+	a''	a'	a'''	a''	a''	+	+	+	a'

**KEY:**

Conventions are as for Table 1, plus the following:

**a'** indicates a sequence of two innovations: first **a**, then a further innovation applying to the output of **a**; **a''** similarly indicates a sequence of three innovations, **a'''** a sequence of four.

... The preconditions for this innovation do not exist in this language.

Abbreviations are:

- 1, 2, 3** first, second, third person
- SG, PL** singular, plural
- E, I** exclusive, inclusive

**SYNTAX**

- 11. 'Leaky' SOV becomes SVO.
- 12. Loss of postpositions (replacement by verb serialisation or zero).
- 13. The possessor shifts from Proto Schouten pre-head position to post-head position:
  - a. where the head NP is alienably possessed;
  - a'. as **a**, plus where the head NP is alienably possessed and the possessor is a pronoun;
  - a''. as **a'**, plus where the head NP is alienably possessed and the possessor is a NP.

**PRONOMINAL PARADIGMS**

- 14. a. The Proto Schouten distinction between 1PL.E and 1PL.I is lost in the subject prefix, object suffix and possessive suffix paradigms, but not in the disjunctive paradigm.
  - a'. As **a**, plus loss of distinction between 1PL.E and 1PL.I in the disjunctive paradigm.
- 15. In paradigm of possessive pronominal suffixes, conflation of forms:

			a.		a'.		a''.	
	SG	PL	SG	PL	SG	PL	SG	PL
1	A	D	1	A	F	1	A	A
2	B	E	2	B	F	2	B	B
3	C	F	3	C	F	3	C	C

a'''. Complete loss of possessive pronominal suffixes.

- 16. a. In paradigm of object pronominal suffixes, replacement of plural forms by disjunctive pronouns:
  - a'. As **a**, plus replacement of 2SG and 3SG forms by disjunctive pronouns.
  - a''. As **a'**, plus replacement of 1SG form by disjunctive pronoun, i.e. loss of all object pronominal suffixes.

**NUMERAL SYSTEM**

- 17. a. Loss of Proto Schouten word for 'ten' (Proto Schouten had 'one' to 'five', and 'ten').
  - a'. As **a**, plus loss of Proto Schouten word for 'five'.
  - a''. As **a'**, plus loss of Proto Schouten words for 'three' and 'four', i.e. reduction to binary system.

## VERB PHRASE MORPHOLOGY

18. Loss of reduplication as marker of continuative aspect (replacement by cliticised aspect markers in some languages).
19. Coalescence of pre-verbal aspect markers with subject pronominal prefixes to form portmanteau subject/aspect prefixes.

## DERIVATIONAL MORPHOLOGY

20. Loss of the nominalising derivational suffix *\*-ŋa*.
21. a. Loss of causative prefix *\*va-*.  
a'. As a, plus loss of reciprocal prefix *\*vai-*.

The pattern which emerges in Tables 1 and 2 is that Manam is the most conservative language of the group,<sup>4</sup> and that, the further one moves away geographically from Manam, the more innovations each language had undergone. However, more detailed patterning is also visible: the languages from Ulau-Suain to Sera form a series of (apparently) four minimal groups, one within the other and each marked off by its own set of shared innovations, shown in Figure 1.<sup>5</sup> The group which includes all five of these languages is the Siau family (Ross 1988:123). On the morphosyntactic evidence of Table 2 (innovations 14a, 16a, and 21a), one could argue for a yet larger fifth group also including Kairiru, but there is no support for this in the phonological evidence of Table 1. We should note in passing that there are no grounds to regard the languages from Medebur to Kairiru (i.e. those not included in the Siau family) as a minimal group, and no strong grounds for recognising any minimal groups among them, other than possibly Kaiep and Kairiru (innovations 6c and 10ic). This means that a traditional genetic tree of the internal relationships of the Schouten languages would look like Figure 2A. However, this fails to capture the nature of the relationship among the Schouten languages, namely that of a linkage (i.e. a group of languages which have arisen by increasing differentiation out of an earlier dialect chain). This is presented in Figure 2B, in accordance with the conventions of Ross (1988:9-11). Figure 2B seeks to capture the following sequence of events:

- (a) Proto Schouten was once a single language;
- (b) Proto Schouten diffused into a dialect chain;
- (c) 'Gaps' in the chain emerged between the Medebur to Wogeo, the Kaiep/Kairiru and the Proto Siau portions of the chain (Proto Siau was apparently a single communalect); these gaps were points at which contact between the speakers on either side was reduced;
- (d) From here,
  - (i) continued differentiation in the Medebur to Wogeo and the Kaiep/Kairiru portions of the chain resulted in the gradual emergence of the Medebur, Manam, Bam, Kis, Wogeo, Kaiep and Kairiru languages;

<sup>4</sup>Phonologically, Bam is more conservative than Manam by one innovation, since it retains the Proto Oceanic/Proto Schouten distinction between *\*n* and *\*ŋ* (Table 1, innovation 8) where Manam has a merger. Morphosyntactically, however, Bam is considerably more innovative than Manam.

<sup>5</sup>Of the four closed groups in Figure 1, the group consisting of Tumleo, Sissano and Sera is more weakly supported than the others, and it could be argued that Tumleo is more closely related to Ali (as shown in Fig. 3 of Ross 1988:122) than to Sissano and Sera. However, this is not important in the present context.

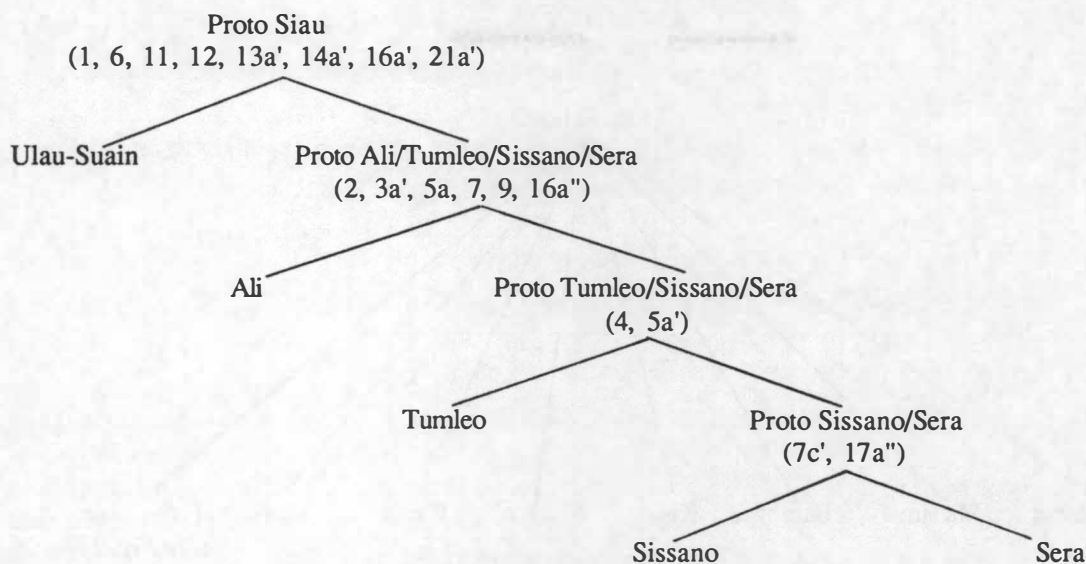


FIGURE 1: THE SIAU FAMILY: A GENETIC TREE  
(Figures in parentheses refer to the innovations in Tables 1 and 2)

- (ii) Proto Siau apparently split first into the language ancestral to Ulau-Suain and then into the language ancestral to Ali, Tumleo, Sissano and Sera; this was followed by the further splits shown in Figure 1.

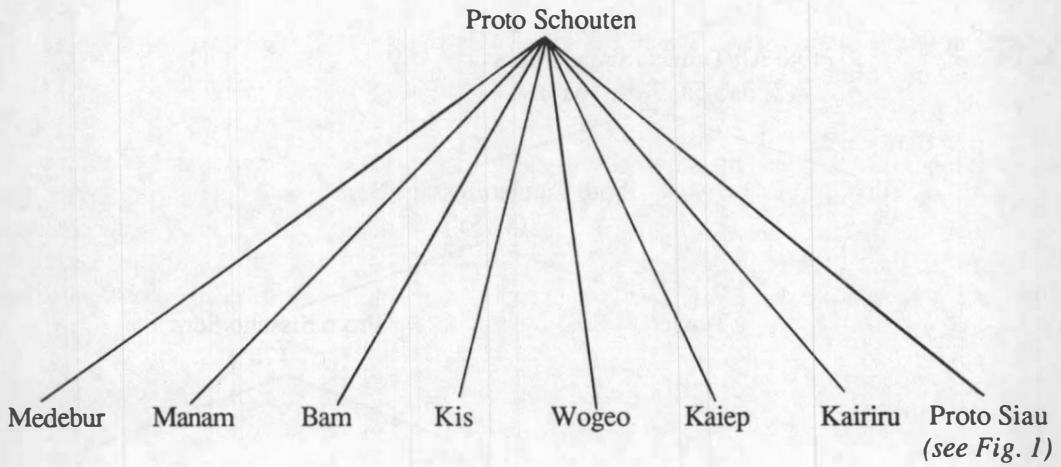
Obviously, Figure 2B provides only an approximate representation of this history, since dialect differentiation and the emergence of 'gaps' in a linkage are in reality a matter of degree, and this gradient is not readily captured in a diagram of this kind.

The cline of innovativeness referred to at the beginning of this paper is clearly visible, then, in the languages of the Schouten chain. Manam, geographically the second language from the eastern end of the chain, is the most conservative, whilst Sera, at the western end of the chain, is the most innovative. In general, the tabulation of innovations in Tables 1 and 2 shows each language from Manam to Sera to be somewhat more innovative than the language immediately to the east of it. The one possible bump in the cline is Kis, which Laycock (pers. comm.), the first linguist to survey the Schouten area, noted was morphosyntactically aberrant, probably because of change induced by one or more of the Papuan (non-Austronesian) languages by which it is surrounded. The sort of change Laycock referred to is exemplified by what has happened to possessive noun phrases in Kis. Proto Schouten, like most Oceanic languages, distinguished between inalienable and alienable possession. In both structures, the possessor noun phrase, or, as in the reconstructed Proto Schouten examples below, the possessor disjunctive pronoun, preceded the possessed noun. In the inalienable structure, as in (1), the possessive pronominal suffix was added directly to the possessed noun; in the alienable structure, it was added to a classifier – in (2) the general classifier \**ne-* – which preceded the possessed noun:<sup>6</sup>

<sup>6</sup>Abbreviations used in examples are:  
CL possessive classifier  
PSR possessor marker



A. 'FAMILY TREE'



B. AN APPROXIMATE REPRESENTATION OF LINKAGES

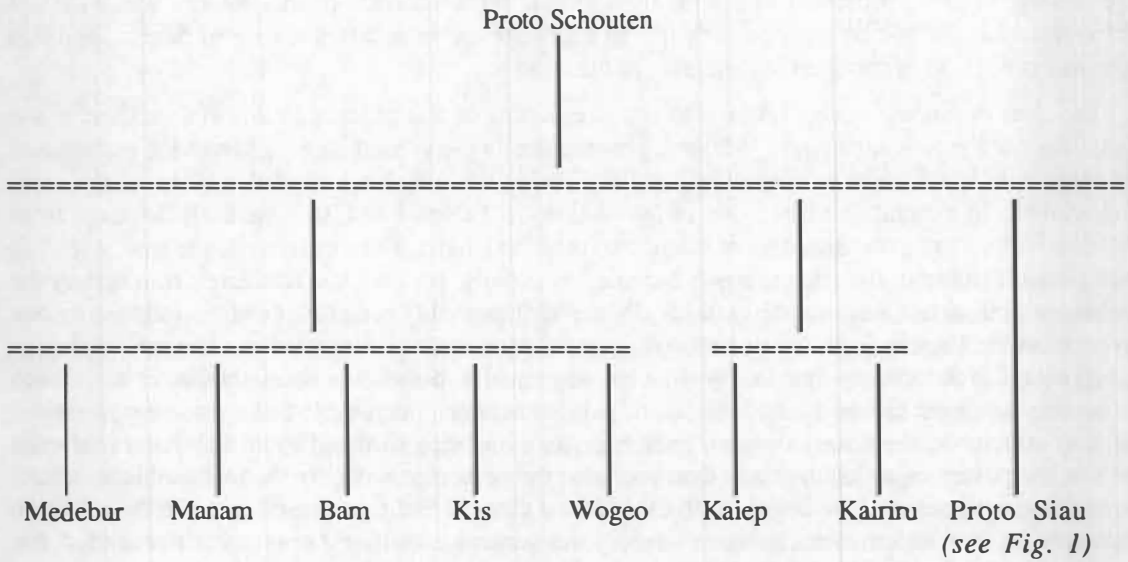


FIGURE 2: THE SCHOUTEN CHAIN: GENETIC TREES

- (1) \**au tama-gu*  
 I father-my  
 my father
- (2) \**au ne-gu ruma*  
 I CL-my house  
 my house

Although a number of Schouten languages have undergone innovation in the possessive noun phrase (see innovations 12 and 14 in Table 2), a radical reanalysis has occurred in Kis. The inalienable/alienable distinction has been lost, as illustrated in (3) and (4) (this has otherwise occurred only in Sera), and the reflex of the general classifier \**ne* – Kis *ni* – has lost its suffixes and been reanalysed as a possessor marker cliticised to the possessor noun phrase or disjunctive pronoun, thus allowing the possessor to shift position without ambiguity, as in (5):

- (3) *au-ni tama*  
 I-PSR father  
 my father
- (4) *au-ni baja*  
 I-PSR house  
 my house
- (5) *baja au-ni*  
 house I-PSR  
 my house

One of Kis's close neighbours is the Papuan language Murik. Kis *ni* functions similarly to Murik *na* (Schmidt 1953), suggesting that the reanalysis of the general classifier as a possessor marker is indeed the result of contact-induced change.

The language which appears to be significantly at odds with the overall pattern of the Schouten chain is Medebur, which is about as innovative as Bam or Wogeo, but is situated in a coastal enclave at the eastern end of the chain. There is substantial evidence from local oral history, however, that the people of Medebur are descended from colonists from Manam, that is, they constitute a relatively recent disturbance to the geography of the cline of innovativeness.

This geographical patterning of innovativeness is curious in that it is not readily explained by the standard canons of historical linguistics. All other things being equal, as the languages of the Schouten chain emerged, be it by differentiation out of a dialect chain or by outright separation, we would expect each language to have undergone innovations of its own. Patterns of shared innovations like those found for the Schouten chain as a whole and in the Siau group are of course probable in cases where two or more languages share a common ancestry. And, since the circulation of Blust (1981), most scholars in historical Austronesian linguistics would probably expect innovation to occur at different speeds in different languages. But this ought to mean that the pattern of innovation across a minimal group of languages is unpredictable and not, as in the Schouten chain, that innovativeness is geographically patterned. Why, for example, are the easternmost languages (Medebur, Manam, Bam) more conservative than their relatives to the west?

If the Schouten chain were the only minimal group in western Melanesia to show this correlation between innovativeness and geography, then it would simply be part of a larger unpredictable

pattern. But this is not what we find. A similar correlation is detectable in a number of groups of varying size (and evidently of varying antiquity) scattered across the region. These include:<sup>7</sup>

- (a) the Kilivila chain (Kilivila, Iwa, Gawa, Muyuw and Budibud), stretching from west to east in a band to the north of the south-east Papuan archipelagos: its most conservative language is the westernmost, Kilivila, with languages increasing in innovativeness as one moves eastwards;
- (b) the Central Papuan family, scattered along the south coast of the Central Province of Papua: its most conservative language is Sinagoro in the east, its most innovative is its westernmost, Mekeo (the small Oumic group is out of geographical sequence, to the east of Sinagoro);
- (c) the Meso-Melanesian cluster, covering east New Britain, New Ireland, Bougainville, and the western Solomons: its most conservative language is the westernmost, Bali-Vitu (off north New Britain) and its most innovative languages are (phonologically) those of Santa Ysabel and (morphosyntactically) those of the Nehan/North Bougainville group, centred on Buka Island ;
- (d) the Lavongai/Nalik family (Nalik, Kara, Tiang, Tigak and Lavongai) of northern New Ireland: its most conservative language is its southernmost, Nalik, its most innovative is its northernmost, Lavongai;
- (e) the Markham family, occupying the Markham Valley and the valleys of several tributaries of the Markham River (cf. Holzkmnecht 1989): its most conservative languages are those of the Busu subfamily (especially Musom and Nafi) at the eastern end of the Markham family, its most innovative those of the Upper Markham subfamily in the west;
- (f) the South Huon Gulf chain, occupying the coast of the Huon Gulf to the south of the Markham River and the mountain ranges and valleys of the hinterland: its most conservative member is Numbami at its (coastal) south-western extreme, its most innovative Mumeng at its (inland) north-eastern extreme.

In these circumstances, the correlation between innovativeness and geography displayed by the Schouten chain cannot be regarded as a chance patterning. I wrote above, "All other things being equal ...we would expect each language to have undergone innovations of its own", resulting in unpredictable patterns of innovation. Clearly, however, "all other things" are not "equal", and it is to these that we now turn.

Apart from historically inherited shared innovations, there are two factors which militate against a pattern of random innovation in individual languages and result in a cross-linguistic patterning of innovations. The first factor is the one which underlies the 'wave' model of historical linguistics, namely that if speakers of related communalects are in contact with each other, it is to be expected that an innovation occurring in one communalect will diffuse to its neighbours. However, the patterning resulting from this diffusion will be a collection of intersecting isoglosses like that shown in (6) below. The pattern in (6) was generated by a crude computer model. In an effort to replicate the parameters in Table 1, the model assumed a chain of twelve communalects and nine possible points of innovation. Fifteen innovations were generated, the model specifying randomly the point of innovation, the communalect in which it first occurred, and whether it spread to neighbouring communalects and, if so, in which direction(s), and the number of communalects on either side of it through which it spread. An aspect of the model's crudity was that the probability of an innovation

<sup>7</sup>The groups referred to here are analysed and mapped in Ross (1988). More detailed information on the Markham group is found in Holzkmnecht (1989).

spreading through, say, eight neighbouring communalects was as high as the probability of it spreading to just one neighbour. As a result, one innovation (at point 2) spread through all twelve communalects. This means that a comparative linguist would (erroneously) attribute it to the proto language and it is therefore not recorded below.

(6)	1	2	3	4	5	6	7	8	9
A	+	-	-	-	-	-	-	-	a
B	+	+	-	-	-	-	-	-	a
C	+	+	+	-	b	-	-	-	a
D	+	+	+	-	ab	-	-	-	ab
E	+	-	+	-	a	-	-	-	b'
F	+	-	+	+	ac	-	-	-	b
G	+	-	+	+	ac	-	-	-	b
H	+	-	+	+	ac	ab	-	-	-
I	+	-	-	+	ac	b'	+	-	-
J	-	-	-	+	c	b	-	-	-
K	-	-	-	+	c	b	-	-	-
L	-	-	-	+	c	b	-	-	-

It is not difficult to perceive from (6) that in this model the communalects at the two ends of the chain are least likely to undergo innovations (because each has only one neighbour from which it can receive a diffused innovation) and that the communalects in the middle of the chain have the greatest probability of undergoing innovations. The 'wave' model, then, does not give rise to a completely unpredictable pattern of innovativeness in a chain: it predicts that the languages at both ends of the chain will (other things being equal) be more conservative than those in the middle. But this provides no explanation of the pattern found in the Schouten chain and the groups listed above, where the languages at one end of the chain are more conservative, but those at the other end more innovative.

The second factor which may result in two or more languages having a similar patterning of innovations is the occurrence of paradigmatically motivated sound change. Most sound changes are syntagmatically motivated, that is, motivated by articulatory or, less often, acoustic considerations relating to their occurrence in the flow of speech (Stampe 1979). Such change appears sometimes to be facilitated by a gap in the system which will accommodate the sound resulting from the change (see, for example, Ross 1988:16-17, 70-71). Sometimes, however, a syntagmatically motivated change adds a new contrast to the paradigm, and the added contrast may result in an unstable paradigmatic structure. A paradigmatically motivated sound change then restores stability to the system. Although a number of scholars have occupied themselves with the implicational relationships which constitute a stable system (e.g. Hockett 1963; Jakobson 1963; Gamkrelidze 1978) and with syntagmatically motivated change (Martinet 1955; Anderson 1973), a general theory of paradigmatic stability and paradigmatically motivated change has to my knowledge yet to emerge, and we are left depending in some measure on ad hoc judgments.

Table 1 shows two reasonably clear cases of paradigmatically motivated change in the Schouten languages, at points 6 and 10. The consonant paradigm of Proto Schouten was as follows:

(7)	<i>pw</i>	<i>p</i>	<i>t</i>	<i>k</i>	<i>q</i>
	<i>bw</i>	<i>b</i>	<i>d</i>	<i>j</i>	<i>g</i>
		<i>v</i>	<i>r</i>		<i>y</i>
			<i>l</i>		
			<i>s</i>		
			<i>z</i>		
	<i>mw</i>	<i>m</i>	<i>n</i>	<i>ñ</i>	<i>ŋ</i>
	<i>w</i>			<i>y</i>	

From the innovations in Table 1, it is evident that there were two points of instability in this paradigm, namely *\*z* (point 6) and *\*q* (point 10). Proto Schouten *\*z* had arisen from a syntagmatically motivated change, namely the lenition of Proto Oceanic *\*s*. Its instability seems to result from two factors. Firstly, *\*s*/*\*z* are the only voiceless/voiced pair outside the stop series. Secondly, *\*z* was almost certainly phonetically [z], that is, alveolar, laminal, voiced and fricative; it was thus rather close in articulation and acoustically to *\*j*, which was probably phonetically [dʒ], that is, post-alveolar, sub-laminal, voiced and affricated. Hence we find at point 6 a number of different strategies, all with the same function – to get rid of *\*z*:

- (8) (a) *\*z* > j, y (merger with *\*j*)  
 (b) *\*z* > r (merger with *\*r*)  
 (c) *\*z* > y, ∅  
 (d) *\*z* > s

The instability of Proto Schouten post-velar *\*q* was similarly the result of two factors: it is the only consonant at its point of articulation, and it was phonetically similar to *\*k*.<sup>8</sup> Proto Oceanic had had two consonants, *\*q* and *\*R*, at a post-velar point of articulation, but in many parts of Oceania the loss of *\*R* (through deletion or merger) resulted in the paradigmatic isolation of *\*q* and its subsequent loss.<sup>9</sup> Here again, we find that the Schouten languages have resorted to a variety of strategies to eliminate the unstable phoneme.

Because the sound changes at points 6 and 10 in Table 1 are motivated by pressure to eliminate a phoneme from the paradigm, they show a patterning which is rather different from that of the rest of the table. Firstly, they occur in *all* the Schouten languages (and are the only changes shown in Tables 1 and 2 for Manam) and, secondly, a variety of strategies occurs to achieve the elimination of each phoneme. Where neighbouring languages share an innovation, this is, of course, attributable either to shared inheritance (in the case of 6d, exclusive to the Siau languages) or to diffusion, but the

<sup>8</sup>I have argued elsewhere that Proto Oceanic *\*q* was probably a glottal, rather than a backed velar, stop (Ross 1988:32). The same was probably true of Proto Schouten *\*q*. In Kairiru we find the odd circumstance that Proto Schouten *\*k* has become /q/ (backed velar stop), whilst *\*q* has become /k/. This virtual reversal of positions of articulation occurred somewhat as follows:

Proto Schouten *\*q*- > *\*y*- > Proto Kaiep/Kairiru *\*g*-  
 Proto Schouten *\*k* > Kairiru *q*  
 Proto Kaiep/Kairiru *\*g*- > Kairiru *k*-

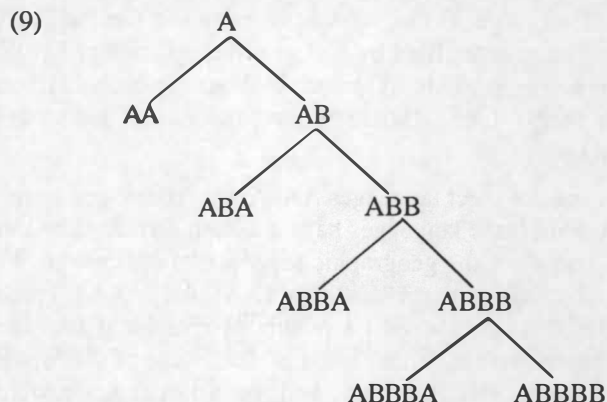
<sup>9</sup>Note that the phonetic similarity of two phonemes alone is apparently not sufficient to trigger paradigmatically motivated change. Buang, for example, is a west Melanesian language which has a velar/backed velar contrast with four manners of articulation (Hooley 1970). The fact that the contrast is thus paradigmatically well supported seems to contribute to its maintenance.



variety revealed in the pattern of innovation across the Schouten languages as a group is diagnostic of paradigmatically motivated changes at these two points.

Again, this patterning of innovations does nothing to explain the conservatism of the languages at the eastern end of the Schouten chain. On the contrary, it provides an explanation of why sound changes *have* occurred at points 6 and 10 in Table 1, and leaves us with a clear correlation between syntagmatically motivated changes (both phonological and morphosyntactic) and geographical location. As I remarked at the beginning of this paper, the most conservative language in a minimal group seems often to be the one which has remained at its present location for the longest period of time, whilst the most innovative is often the one which has moved furthest from the 'home' location of the group. It is to the arguments for this that I now turn, continuing to use the Schouten chain as an illustration.

Whilst linguistic evidence alone never provides direct information about the home location of the proto-language of a group, it sometimes indicates – through the shape of the group's genetic tree – the probable course of migration, and this often gives a clue as to the home location. In Figure 1, the genetic tree of the Siau family has an asymmetrical recursive branching structure like that in (9):



Each node represents the proto-language of a minimal group, and the simplest interpretation of the asymmetry of the tree is in terms of Grace's (1986:4) characterisation of Oceanic migrations: "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". On this interpretation, the 'A' branch from each node represents a 'stay-at-home' group, whilst the 'B' branch represents the group which has moved out beyond the present periphery. Thus the speakers of language AA are descended from the earliest group of stay-at-homes. This interpretation also corresponds with Dyen's (1956:624) principle that "if non-co-ordinate units are contiguous, and neither of them nor any of their respective unshared proto-languages is a migratory unit, their areas are probably the area of their shared proto-language". That is, if languages AA and ABA are geographic neighbours, their location is probably the home location. On this interpretation of Figure 1, the homeland of Proto Siau was probably somewhere in or around the area occupied by today's Ulau-Suain-speaking villages.

Genetic trees of the shape shown in (9) crop up either in this or in a modified but recognisable form at several points in western Melanesia. Cases in point are the Mengen family of south-east New



Britain (Ross 1988:162), the South Huon Gulf family,<sup>10</sup> and the Lavongai/Nalik family of northern New Ireland.<sup>11</sup> The genetic tree of the Austronesian family as a whole, as reconstructed by Blust (1985) also has this shape: AA is the Formosan languages, AB Proto Malayo-Polynesian, ABA the Western Malayo-Polynesian languages, ABB Proto Central/Eastern Malayo-Polynesian, ABBA the Central Malayo-Polynesian languages, ABBB Proto Eastern Malayo-Polynesian, ABBSA the South Halmahera/West New Guinea group and ABBSB Proto Oceanic. This implies the conclusion that Taiwan is the homeland of Proto Austronesian, a conclusion generally accepted by scholars (Blust 1985; Bellwood 1985:110-111). Similarly, the Central Pacific genetic tree (cf. Clark 1973:18) has the approximate form of (9), where AA is the Fijian dialects, AB Proto Polynesian, ABA the Tongic languages, ABB Proto Nuclear Polynesian, ABBA the Samoic-Outlier languages, and ABBS Proto Eastern Polynesian. The implied conclusion is that Proto Central Pacific was spoken somewhere in the Fijian archipelago and that the ancestors of the Proto Polynesian speakers left there and settled in the Tongan archipelago. Again this accords with scholarly opinion (Geraghty 1983).

Frequently, of course, the genetic trees of language groups in western Melanesia have a shape which is either a heavily modified version of (9) or which is unrecognisable as (9). There are basically two kinds of modification. The first is simply that the regular asymmetry is disturbed when an 'A' branch divides and becomes a group of languages (as in the Austronesian and Central Pacific families). The second occurs when languages have diversified by dialect differentiation rather than by separation. The resulting dialect chain does not provide information about the probable home location in the way that the tree structure in (9) does. Hence Figure 2 does not assist us in determining the homeland of the Schouten group.

For genetic trees with the shape of (9), we would expect languages AA, ABA, ABBA and so on to reflect the sequence in which settlement occurred. If the languages have not been displaced by more recent events, their present locations should also show the geographic sequence of settlement. This inference is supported for the Siau languages by the fact that Ulu-Suain (AA), Ali (ABA), Tumleo (ABBA), Sissano (ABBSA) and Sera (ABBSB) do indeed lie in this geographical sequence. Among the other groups noted as reflecting the structure in (9), the same is clearly true of the Mengen family, the South Huon Gulf family, the Lavongai/Nalik family, and, on a larger scale, of the Austronesian family as a whole and of the Central Pacific family.

The significant point here is not simply that the shape of the genetic tree indicates which present-day language is nearest to the likely home location of the proto-language: it is that in each of the cases where there is a correlation of the tree-shape with geographical distribution, language AA is the most conservative of the group. We have already seen that this is true of Ulu-Suain in regard to the Siau group. It is also true of Uvol in the Mengen family, Numbami in the South Huon Gulf family, Nalik in the Lavongai/Nalik family, the Formosan languages in the Austronesian family, and Fijian in the Central Pacific family. In a recent paper, Blust (1990) has observed such a correlation for the Austronesian language family as a whole.

Other kinds of evidence which help to establish the home location of a proto-language are non-linguistic. Grace (1986:5) proposed the "principle of shortest moves", to the effect that 'in the

<sup>10</sup>In terms of the tree structure in (9), AA = Numbami, ABA = Kaiwa, ABBA = Hote, and ABBS = the Buang chain (cf. Ross 1988: 154-160). In Ross (1988), the South Huon Gulf family is called the 'South Huon Gulf chain': this is erroneous in relation to the terminology of that work.

<sup>11</sup>In terms of the tree structure in (9), AA = Nalik, ABA = Kara, ABBA = Tiang, ABBSA = Tigak and ABBSB = Lavongai (cf. Ross 1988:291).

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'. Although he was proposing it as a methodological principle (which would provide working hypotheses), its application works well in practice for much of Oceania. Given that Proto Schouten is a part of the North New Guinea cluster, the dialect chain ancestral to which was located in the region of the Vitiaz Strait, well to the east of the Schouten group, it is a reasonable inference that settlement of the Schouten area took place initially from east to west. This in turn implies that the most likely home location for Proto Schouten is the location of the Schouten language closest to the Vitiaz Strait, that is, furthest to the east. If we set aside Medebur for the reason mentioned above, we may infer that Manam Island is the most probable home location of Proto Schouten. Since the genetic tree has allowed us to establish the location of Proto Siau at the eastern end of the Siau minimal group, indicating that the Siau area was settled from east to west, the inference that most of the area occupied by the Schouten group was settled in that direction from Manam comes as no surprise.

Two other kinds of evidence which can help ascertain the home location of a proto-language are oral history and archaeology. While it is clearly not conclusive, oral evidence that the ancestors of both Medebur and Kis speakers came from Manam supports the inference that it was the home location.<sup>12</sup> There are unfortunately no published archaeological findings which might relate to Schouten languages. In one significant case, however, we do have archaeological corroboration of a homeland hypothesis: a consensus has emerged among archaeologists studying the Lapita culture that this culture had a major centre in the neighbourhood of Talasea (Willaumez Peninsula, north New Britain), whence obsidian was obtained and distributed, and that the Lapita potters are to be identified with the speakers of Proto Oceanic (Spriggs 1984; Bellwood 1985:125-127; Allen and White 1989). There is also archaeological support for the claim that the Fijian archipelago was the homeland of Proto Central Pacific.

I have set out above the methods at our disposal for inferring the most probable home location of a proto-language, and also indicated several cases in which the language currently spoken at the putative home location is the most conservative in its minimal group. I have also offered an interpretation of innovations which have occurred in the Schouten languages and shown that the accepted paradigm of historical linguistics offers no explanation for the observed correlation between the cline of innovativeness and geography. It does not seem to me particularly radical to suggest that sedentariness is related in a principled way to conservatism and provides the missing explanation.

Grace (1985:11) points out that if languages change at a variable rate, it is probably necessary to look to causes external to language for an explanation of this variability. The correlations we have noted in this paper suggest that when a community migrates from one location to another, this is likely to result in innovation. Grace (1985:10) suggests that "languages in a state of accelerated change would be characterised by a relatively high degree of permissiveness as regards standards of correctness". It is not difficult, a priori, to conceive how such permissiveness might arise. If Oceanic settlement occurred through small groups who left their home villages on the periphery of the region in which Oceanic was spoken and moved to a place beyond that periphery, then it is probable that the settlers were usually led not by the gerontocrats of the village, but by the middle generation of men with young families. The new settlement would then have a different generational and authority structure from the old. As a result the speech patterns of the younger generation, which tend in most cultures to be innovative, would form a larger proportion of daily discourse than before, and, in the

<sup>12</sup>The oral evidence from Medebur was collected by the writer, from Kis by Laycock (1976:82).

absence of the former elders would go unchecked. Change resulting from this situation would accord with the probably innovative values of the new migrants, and, as the new community developed its own sense of identity, more novel forms of speech would become emblematic of communal identity and thereby entrenched. Since in many cases the new community would have maintained contact with the old, members of the old community would be aware of this linguistic innovation, and would emphasise the conservative features of their own speech as emblems of *their* identity.

In the context of western Melanesia, of course, migration to a new location often entailed contact with speakers of previously unknown alien languages. In some cases, this resulted in bilingualism which in turn led to contact-induced change in the Oceanic language of the migrants (Laycock 1979; Ross 1987). Whilst this undoubtedly was a factor in language change, however, we need to view it in the context of the values of the community. Tiesler (1969-1970) shows on the basis of ethnographic accounts that before European contact, the most regular contacts of the speakers of Bam, Wogeo, Kaiep, Kairiru, Ulau-Suain, Ali, Tumleo, Sissano and Sera were with each other, not with their non-Oceanic speaking neighbours. The two groups which were more frequently in contact with non-Oceanic speakers than with speakers of other Schouten languages were the Manam and the Kis. As we have seen, the two groups vary considerably in innovativeness. As was noted above, Kis clearly shows in its morphosyntax the marks of contact-induced change, whereas Manam shows no sign thereof. This difference can be attributed to two factors, which are not mutually exclusive. Firstly, the Manam speech community is far larger and geographically more isolated than Kis, which could, but need not, put a brake on contact-induced change.<sup>13</sup> And secondly, for the reasons discussed in the previous paragraph, contact-induced change is likely to have been far more readily accepted by the Kis than by the Manam.

A methodological consequence of recognising a correlation between sedentariness and conservatism is that, wisely used, it can help in the interpretation of linguistic prehistory, especially in those cases where (i) the evidence of conservatism, geographic distribution and genetic-tree shape is at odds with the principle of shortest moves, or (ii) the genetic-tree shape and/or geographic distribution is too confused to allow inferences about migration.

A case of the former kind is provided by the South Huon Gulf family, mentioned above. Such is the location of the family (for a map, see Ross 1988:134) in relation to the Proto Oceanic homeland that we would expect settlement to have taken place along the coast from north to south. But the correlation of innovativeness and genetic-tree shape suggests the opposite: Numbami, the most conservative language of the group, is the southernmost, and there is a progression of innovations northwards, first along the coast and then inland. As a result, we must seriously entertain the hypothesis that the principle of shortest moves does not apply, and that the ancestors of the speakers of Proto South Huon Gulf in fact made one relatively long move (perhaps from the Huon Peninsula straight across the Huon Gulf) to the area where Numbami is situated today.

A case in which both genetic-tree shape and geographic distribution are too confused to allow inferences about migration is the Ngero/Vitiaz family.<sup>14</sup> The branches of this family, together with the languages which are relevant here, are shown in Figure 3. As the map indicates, these languages

<sup>13</sup>Laycock (1973) gives the 1970 population of Manam as 5070 and of Kis as 216. The example of contact-induced change discussed by Ross (1987) concerns an even larger offshore island population than Manam, so it cannot be taken for granted that the present size of the speech community has any causal relationship with contact-induced change.

lie in a confused scattering through the Vitiaz Strait area. Because Figure 3 lacks the neat asymmetry of (9) above, we cannot tell with certainty whether it is Proto Ngero or Proto Vitiaz which occupies the position of language AA, that is, the language which is probably nearest the homeland. However, the fact that the Ngero branch of the family has undergone far less diversification than the Vitiaz branch implies that the Proto Ngero was the original stay-at-home language, although the furthest members of the Bariai branch (Kove and Malalamai) are today about two hundred kilometres apart. This implication is supported by the conservatism of the Ngero languages in comparison with the Vitiaz languages. Malalamai is so similar to the other members of the Bariai branch that we must assume it to have reached its present location quite recently. Within the Ngero family, the Bariai branch, mainly on New Britain, is the more conservative: its languages are among the few Western Oceanic languages not to have completely merged Proto Oceanic \*r and \*R (Ross 1988:175-177, 387-388). Furthermore, the principle of shortest moves from the Proto Oceanic homeland suggests that the homeland location of Proto Ngero, and therefore putatively of Proto Ngero/Vitiaz, was on New Britain, and archaeological evidence indicates that the Tuam languages have been in their present location for only about 300 years (Lilley 1988). Within the Vitiaz linkage, the most conservative language appears to be Mangap, located on Umboi Island, and it is a reasonable hypothesis that Proto Vitiaz was located on this, the largest island in the Strait.

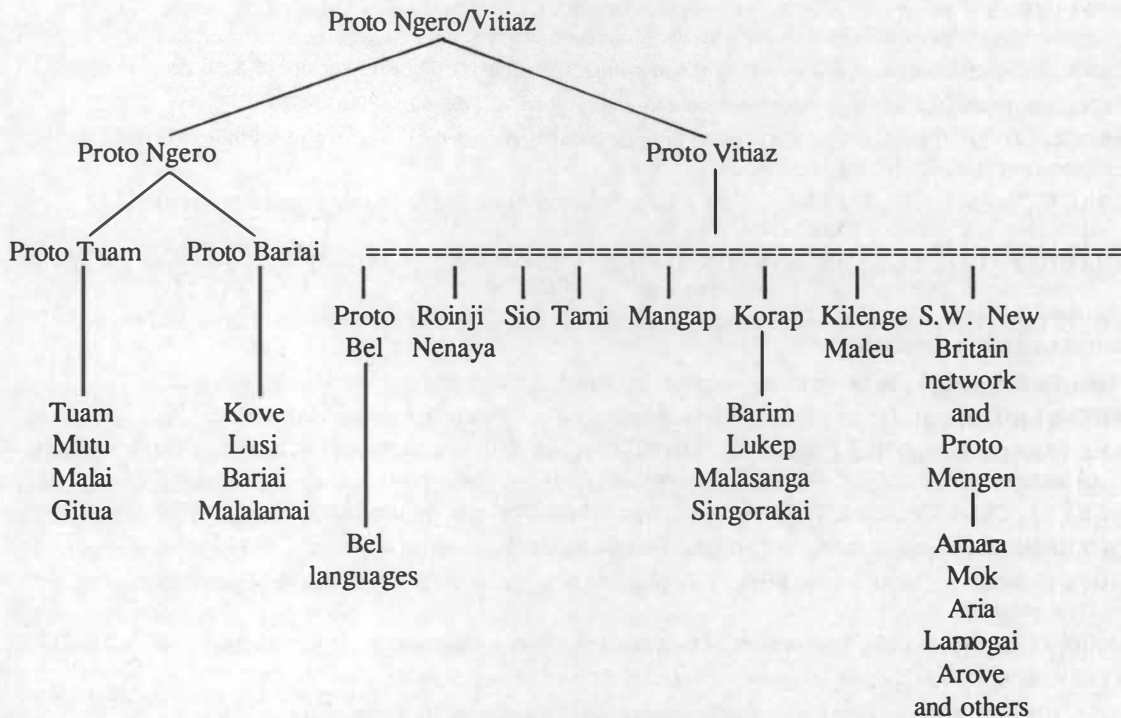


FIGURE 3: NGERO/VITIAZ FAMILY: GENETIC TREE

<sup>14</sup>See Ross (1988:160-183) for further detail on the Ngero/Vitiaz family. I would like to acknowledge here the role of Lincoln (1977) in laying the groundwork of my understanding of the Ngero and Korap groups (which he named).

Finally, but certainly not conclusively, I would like to make two comments. The first is that, as I hope the last example indicates, conservatism cannot be used as an automatic discovery procedure: it is simply one factor which needs to be taken into account in conjunction with the others discussed here. The second comment is that, as will be very clear to some readers, my position on the interpretation of linguistic prehistory is different from that of scholars such as Bailey (1980) and Thurston (1987). But that is not a matter to be handled in a concluding sentence.

## BIBLIOGRAPHY

- ALLEN, Jim and J. Peter WHITE, 1989, The Lapita homeland: some new data and an interpretation. *Journal of the Polynesian Society* 98:129-146.
- ANDERSON, James M., 1973, *Structural aspects of language change*. London: Longman.
- BAILEY, Charles-James N., 1980, Old and new views on language history and language relationships. In Lüdtke, ed. 1980:139-183.
- BELLWOOD Peter, 1985, *Prehistory of the Indo-Malaysian archipelago*. Sydney: Academic Press.
- BLUST, Robert A., 1981, Variation in retention rate among Austronesian languages. Paper presented at the Third International Conference on Austronesian Linguistics, Bali, Indonesia.
- 1985, The Austronesian homeland: a linguistic perspective. *Asian Perspectives* 26:45-67.
- 1990, Patterns of sound change in the Austronesian languages. In Philip Baldi, ed. *Linguistic change and reconstruction methodology*, 231-267. Berlin: Mouton de Gruyter.
- CLARK, Ross, 1973, *Aspects of Proto-Polynesian syntax*. Auckland: Linguistic Society of New Zealand.
- DYEN, Isidore, 1956, Language distribution and migration theory. *Language* 32:611-626.
- GAMKRELIDZE, Thomas V., 1978, On the correlation of stops and fricatives in a phonological system. In Greenberg et al, eds 1978:9-46.
- GERAGHTY, Paul, 1983, *The history of the Fijian languages*. *Oceanic Linguistics* special publication 19. Honolulu: University of Hawaii Press.
- GERAGHTY, Paul, Lois CARRINGTON and S.A. WURM, eds, 1986, *FOCAL II: papers from the Fourth International Conference on Austronesian Linguistics*. PL, C-94.
- GRACE, George W., 1985, Oceanic subgrouping: retrospect and prospect. In Pawley and Carrington, eds 1985:1-18.
- 1986, Further thoughts on Oceanic subgrouping. In Geraghty, Carrington and Wurm, eds 1986:1-12.
- GREENBERG, Joseph H., ed., 1963, *Universals of language*. Cambridge, Mass.: MIT Press.
- GREENBERG, Joseph H., Charles A.FERGUSON and Edith A.MORAVCSIK, eds, 1978, *Universals of human language*, vol.2: *Phonology*. Stanford: Stanford University Press.
- HOCKETT, Charles F., 1963, The problem of universals in language. In Greenberg, ed. 1963:1-22.
- HOLZKNECHT, Susanne, 1989, *The Markham languages of Papua New Guinea*. PL, C-115.
- HOOLEY, Bruce A., 1970, Mapos Buang – Territory of New Guinea. Ph.D. dissertation, University of Pennsylvania.
- JAKOBSON, Roman, 1963, Implications of language universals for linguistics. In Greenberg, ed. 1963:208-219.
- LAYCOCK, D.C., 1973, *Sepik languages: checklist and preliminary classification*. PL, B-25.
- 1976, Historical research in Austronesian languages: Sepik Provinces. In Wurm ed. 1976:73-93.
- 1979, Multilingualism: linguistic boundaries and unsolved problems in Papua New Guinea. In Wurm, ed. 1979:81-99.
- LAYCOCK, D.C. and W. WINTER, eds, 1987, *A world of language: papers presented to Professor S.A. Wurm on his 65th birthday*. PL, C-100.
- LILLEY, Ian, 1988, Prehistoric exchange across the Vitiaz Strait, Papua New Guinea. *Current Anthropology* 29:513-516.



- LINCOLN, Peter C., 1977, Subgrouping across a syntactic isogloss. Mimeo.
- LÜDTKE, Helmut, ed., 1980, *Kommunikationstheoretische Grundlagen des Sprachwandels*. Berlin: Walter de Gruyter.
- MARTINET, André, 1955, *Economie des changements phonétiques*. Berne: Francke.
- PAWLEY, Andrew K. and Lois CARRINGTON, eds, 1985, *Austronesian linguistics at the 15th Pacific Science Congress*. PL, C-88.
- ROSS, Malcolm D., 1987, A contact-induced morphosyntactic change in the Bel languages of Papua New Guinea. In Laycock and Winter, eds 1987:583-601.
- 1988, *Proto Oceanic and the Austronesian languages of western Melanesia*. PL C-98.
- SCHMIDT, Joseph, 1953, Vokabular und Grammatik der Murik-Sprache in Nordost-Neuguinea. *Micro-Bibliotheca Anthropos* 3.
- SPRIGGS, Matthew, 1984, The Lapita Cultural Complex. *Journal of Pacific History* 19:202-223.
- STAMPE, David, 1979, *A dissertation on natural phonology*. Bloomington: Indiana University Linguistics Club.
- THURSTON, William R., 1987, *Processes of change in the languages of north-western New Britain*. PL, B-99.
- TIESLER, Frank, 1969-70, Die intertribalen Beziehungen an der Nordküste Neuguineas im Gebiet der kleinen Schouten-Inseln. *Abhandlungen und Berichte des Staatlichen Museums für Völkerkunde Dresden* 30:1-122, 31:111-195.
- WURM, S.A., ed., 1976, *New Guinea area languages and language study*, vol.2: *Austronesian languages*. PL, C-39.





## WILLIAM ANDERSON'S HAWAIIAN WORD LIST

ALBERT J. SCHÜTZ

### 1. INTRODUCTION<sup>1</sup>

As the exploration of the South Pacific revealed an exotic world to the artists and scientists on Captain Cook's three expeditions, there was often no choice for them but to perceive the new in terms of the old. In his thought-provoking study *European Vision and the South Pacific*, Bernard Smith (1984) noted that the expeditions' artists had to interpret visual images according to the "pictorial conventions of the time". Not only was their perception filtered through their own culture, but often it was further altered – sometimes subtly, sometimes blatantly – by the imagination of the engraver or painter who prepared the final version for the public.

In Smith's work, the depiction of tropical flora provides dramatic evidence of the differences between the original drawing and the final product. For example, in a view of Omai's return to Tahiti,<sup>2</sup> the engraver's version of the background vegetation looks rather more like spiders on stalks, or a presage of Wyndham's truffids, than the palm trees that were intended.

But it was not only the pictorial records that suffered in the transfiguration of primary records. In a similar but less obvious way, the linguistic data from Cook's third voyage also changed as they passed from manuscript to print. I should like to discuss one example: William Anderson's Hawaiian word list, which provided the western world with its first extended glimpse of the Hawaiian language.

Published in 1784, sold out in three days, and in its third edition by the next year (Beaglehole 1974:692), the account of Cook's third voyage gave an eager and impatient English-reading public the final instalment of one of the world's greatest adventure stories. The readership became even greater within the next few years with the appearance of French and German translations.

---

<sup>1</sup>I should like to express my appreciation to the staff of the Hawaiian and Pacific Collections, Hamilton Library, University of Hawai'i for the competent assistance which has long been their trademark. For help with stylistic matters, I am grateful to Ann M. Peters and Kenneth L. Rehg.

Keeping George W. Grace unaware of the work on this paper has been rather a hardship. Normally – that is, as I have done for the nearly thirty years that we have been colleagues – I would have asked for his help and advice, which he has always been so generous in giving. Just this once, however, without hearing any outlines or reading any rough drafts, perhaps he can read straight through to the denouement without the usual sense of familiarity with the material.

<sup>2</sup>Smith (1984:115), reproduced from Rickman's (unofficial) edition of Cook's journal from his third voyage (1781).

Aside from Cook's own transcription of a few Hawaiian words in an English context, most of the language sample was collected by Anderson, surgeon aboard the *Resolution* and unofficial naturalist for the expedition. The list appears in the Appendix to the three-volume work: 229 words and phrases, plus another implicit 10, the numerals "as at Otaheite".

At the time of the first European contact with Hawaiian, there was no established method for writing a previously unwritten language. Thus, the most common way to create an orthography for a new language was to use one's own, illustrating unfamiliar sounds with the closest (in this case) English equivalents. Naturally this practice distanced readers from the foreign language, to a degree depending on the accidental fit between it and the compiler's own. In addition, the compiler's intelligence and linguistic sophistication also had a marked effect on the success of the resultant alphabet.

Cook himself did not always receive high marks for his rendering of Hawaiian words. Perhaps the most negative assessment of his Polynesian spelling was that it was a "rough, inconsistent, quasi-phonetic spelling in Latin characters" (Wise and Hervey 1952:311).<sup>3</sup> As for Anderson's spelling of Hawaiian, it has been – at best – superficially treated, or – at worst – ignored.

One of the possible reasons that Anderson's word list has not been taken seriously is that through an unfortunate transition from the original to the final version, it apparently suffered a fate similar to that of the artist's rendering of the Tahitian palm trees.

## 2. ANDERSON'S ANONYMITY

One might propose that the role of a historian can be compared to that of a director or a stage manager. Even if the plot and cast of characters can seldom be altered, the focus can. Compared with the rest of Cook's officers and scientists, Anderson was seldom in the centre of the narration, outshone by the brilliant if harsh light of such characters as the difficult Johann Reinhold Forster.<sup>4</sup> Perhaps it was because of the implicit competition from the natural scientists, for although they all collected samples for the European specimen case, zoology and botany were more showy. Indeed, it is mostly through botany that Anderson's name is known at all, for the genus *Andersonia* serves as a memorial to the man and his work.

At any rate, Anderson's linguistic work has been largely neglected for the past two centuries. For example, the writers of a study of the development of the Hawaiian spelling system (Wise and Hervey 1952) discussed Cook's own spelling of the Hawaiian words that were scattered through his description, and that of several of his officers, but failed to mention Anderson or his word list.<sup>5</sup> A later and mainly derivative article (Walch 1967) made the same omission. When his orthography was finally treated at length (Hervey 1968), it was almost totally misinterpreted.

Even the particulars of William Anderson's life are difficult to find. The *Dictionary of national biography* is cursory in its treatment, giving only the date and cause of his death, his official positions on Cook's second and third voyages, and a summary of his contribution as a naturalist.

<sup>3</sup>One wonders, of course, how it could have been otherwise. Besides, the authors' examples (p. 311) show a greater degree of consistency than they either noticed or were willing to admit.

<sup>4</sup>For example, a recent three-part television movie on Cook's voyages, which was shown in Honolulu in October and December 1989, obliterated Anderson's role entirely.

<sup>5</sup>Perhaps the reason is revealed in the references, in which we find only secondary sources, with no mention of the official publication of Cook's journals.

Mostly through the extensive research of J. C. Beaglehole, we can add some detail to the rough sketches of the official biographies. William Anderson was born in Scotland in 1750 and was educated at Edinburgh University. In December 1771 he joined Cook's crew for the second voyage, and the two ships – the *Resolution* and the *Adventure* – set sail from Plymouth on 13 July 1772. Anderson's position aboard the *Resolution* was that of surgeon's mate. For the third voyage, he was surgeon on the same vessel. Beaglehole (1974:500), summarising Anderson's abilities and contributions to the expedition, wrote that he “was clearly one of the best minds of all the three voyages – professionally competent, but with an interest in all the departments of natural history as they were known at that time, acute as well as wide-ranging, and with a linguistic talent both eager and careful.”

Circumstances almost prevented Anderson from seeing Hawai'i at all.<sup>6</sup> When the expedition was ready to sail north from Tahiti toward the American north-west coast, Anderson and Charles Clerke (Captain of the *Discovery*) had serious reservations about leaving the tropics. Because they were both ill with tuberculosis (or consumption, to give the disease its contemporary name), they did not wish to “encounter the severities of a frozen climate”, but instead preferred to stay in the Society Islands. However, as the ship progressed from one island to another, it was always the next one that would be the place to remain. Why the two men continued the journey north is still a minor mystery; Beaglehole suggested that it was Clerke's sense of duty that lay behind his procrastination with the paper work that would have allowed him to resign his command. Perhaps we can hint further that Anderson might have dared the adventure with a companion, but hesitated to risk it alone.

### 3. ANCHORAGE AT WAIMEA, THE VISIT TO A *HEIAU*, AND THE COLLECTION OF THE WORD LIST

On 18 January 1778, as Cook and his crew were bound north-west from Christmas Island, they sighted land again – the Hawaiian Islands. They advanced toward the three northern islands and arrived at the eastern end of Kaua'i on the afternoon of the 19th. They then proceeded to the west, eventually anchoring in Waimea Bay, off the village and mouth of the river by the same name.

Just before landing at Kaua'i, Anderson had been very ill. But on 21 January 1778, he was sufficiently recovered to walk with Cook and Webber to a *heiau*<sup>7</sup> about a mile up the river (Beaglehole 1974:574). As we can surmise from his journal, Cook was carefully noting nearly all that went on around him on the way to and at the *heiau* (Beaglehole 1974:576). Webber spent the time sketching, resulting in a series of drawings;<sup>8</sup> Cook noted this as well. But he did not report on Anderson's activities. It is obvious from his word list, however, that he was recording, in his own way, the same scene that Cook observed and Webber drew, for part of his word list is an inventory of many of the ceremonial items in the *heiau*.

<sup>6</sup>This discussion is summarised from Beaglehole (1974:568-569).

<sup>7</sup>It is not certain which of several *heiau* sites up the Waimea River the men visited, but from Samwell's journal (Beaglehole 1967:1083), we learn that it was about a mile upstream and “on the Banks of the River”. Several sources specify that it was Ke'a'ali'i *heiau*; Bennett (1931:104) wrote that “so little remains that the rumor can not be substantiated”.

<sup>8</sup>Copies of the sketches can be found in Joppien and Smith's descriptive catalogue (1987:418-421).

## 3.1 TAHITIAN AS A LINGUA FRANCA

In contrast to many of the expedition's linguistic encounters, communication in Hawai'i seems to have posed few difficulties. It was a familiarity with Tahitian that eased the way towards a rapid understanding between Cook's crew and the Hawaiians – at least at a superficial level. As a matter of fact, the similarity of Hawaiian and Tahitian was one of the first things that several journal writers mentioned, with a touch of surprise because of the great distance from Tahiti. James King wrote: :

... what more than all surpris'd us, was, our catching the Sound of Otaheite words in their speech, & on asking them for hogs, breadfruit, yams, in that Dialect, we found we were understood ...

(Beaglehole 1967:264n.)

David Samwell was even stronger in his opinion, writing that the language was “the same as that of Otaheite”.

Thus, since Anderson was already familiar with ‘the idiom’, there was little need to use gestures or ad hoc sign language to gather information. For example, it was undoubtedly Tahitian (and perhaps Māori) that helped him elicit the Hawaiian equivalents of three rather difficult grammatical concepts that begin the list: the Hawaiian forms for ‘where (specifically)?’, ‘where (generally)?’, and ‘no’. From there on, the list continues with mostly content words: names for body parts, plants and food, and the artefacts mentioned earlier.

Unless Anderson had further contact with Hawaiians on board the *Resolution*, the time spent at the *heiau* was the extent of his work with native speakers. Storms and heavy seas curtailed repeated trips ashore, and the expedition left the islands on 2 February and headed north, Cook himself having gone ashore only three times. Anderson was able to write in his journal until 3 June 1778 (Beaglehole 1967:cxc), but died exactly two months later at the age of twenty-eight.<sup>9</sup> King wrote: “If we except our Commander, he is the greatest publick loss the Voyage could have sustained” (Beaglehole 1974:614n.). No memorial marks his resting place, for he was buried at sea. The next day, Cook named an island for him (in the Bering Sea), but unfortunately, “Anderson Island” was already St. Lawrence Island, sighted and named by Bering fifty years earlier.

## 4. THE PARADOX

Although Anderson was described as “an extremely intelligent person” (Beaglehole 1974:299), “with a linguistic talent both eager and careful” (1974:500), in the first few decades after Cook, visitors to Hawai'i complained that early word lists were of little help in trying to communicate with the Hawaiians.<sup>10</sup> For example, de Freycinet reported in 1819 (1978:152) that “... the vocabularies of the Sandwich tongue we had on board were so defective, and the spelling so little adjusted to our way of pronouncing, that it was almost impossible to make ourselves understood except by signs.”

To those of us familiar with modern written Hawaiian, one of the first things that strikes us about Anderson's list is that *t* and *r* are written regularly, to the exclusion of *k* and *l*. However, although information on even the current distribution of [t] is surprisingly scanty,<sup>11</sup> it is clear from the

<sup>9</sup>Clerke died on 22 August 1779. He was only thirty-eight: just ten years older than Anderson.

<sup>10</sup>Although de Freycinet did not refer to Anderson's list specifically, one might assume that familiarity with the Cook volumes was de rigueur for any explorer bound for the South Pacific.

<sup>11</sup>Newbrand (1951) collected data from two informants from the Hanalei side of Kaua'i, both with [k] exclusively. However, an 83-year-old (in 1950 or so) man from Waimea, Kaua'i, produced [t] not in his narrative, but under special



missionary records of the 1820s that [t] was common, if not universal, on both Ni'ihau and Kaua'i, and some other parts of the Hawaiian Group as well. As for what Anderson wrote as *r*, the same records also contain sophisticated phonetic descriptions of the sound, showing that what was heard at that time was most likely an alveolar flap [r], similar to /r/ in Tahitian and Māori today.

Other differences from modern spelling, however, are not so easy to explain. When we examine Anderson's orthography carefully, it does seem to leave much to be desired. For example, the common word *hele* 'go' was written as *haire* (and *pele* as *paire*); *he i'a* 'fish' as *haieea*, *he niu* 'coconut' as *eeneoo*, *au* 'I' as *ou*, and *he ihu* 'nose' as *eehehu*, with the unusual sequence of three *es*. *Oo* seems indiscriminately used for both /o/ and /u/, and there are many instances of *y* used as a vowel.

If one took the list at face value – that is, on its own and out of context – an analysis like the following might result. Hervey (1968:24) wrote that although Anderson gave no guide to the orthography he used, it was possible to reconstruct one. He went on to list such a system, one that bears little resemblance to Hawaiian. For example, his list of consonant phonemes includes /b/, /d/ and /f/, none of which is part of the system. As for the vowels, Hervey seemed to assume that all the vowel letters had their current values, except *oo*, which he correctly interpreted as /u/. Even fairly consistent relationships, such as Anderson's use of *ai* for /e/,<sup>12</sup> were explained as sound changes.<sup>13</sup>

Tables, included to show a variety of Anderson's spelling conventions, reflect little basic knowledge of Hawaiian grammar and linguistic history. For example, *heraee* [he lae] 'forehead' was annotated as follows: "Nothing resembling 'heraee' was listed in Pukui & Elbert [1957]. 'He' was most likely the demonstrative used at the beginning of the phrase." The explanation of *matta* [maka] 'eye' reveals that Hervey was ignorant of the extent of the [t] pronunciation on Ni'ihau and Kaua'i. He continued: "It is just possible that a shift from [t] to [k] did not constitute a phonemic change."<sup>14</sup> A fairly obvious form, *Heoo* 'nipple' ([he ū] 'breast'), was interpreted as *hiu*, *hi'u*, *heo*, or *heu* – but not *he ū*.

## 5. THE TAHITIAN PRONUNCIATION KEY

Even with an understanding of Hawaiian phonology and a rudimentary knowledge of Hawaiian grammar, one still finds problems with Anderson's transcription. However, most of this confusion vanishes when the Hawaiian list is examined not in isolation, but as a part of Anderson's total work, and with the knowledge that he was unable to advise the editor or printer about the conventions he used.

Although Hervey was correct in saying that there is no guide to Anderson's orthographic conventions for the Hawaiian list specifically, there is certainly such a guide to his Tahitian list (Cook

---

circumstances: "The recollection of old days brought forth the exclamation 'maita'i! maita'i!' (good! good!) – showing the use of the [t] allophone of /k/" (p. 79). Elbert and Pukui (1979) discussed the use of [t] on Ni'ihau, and mentioned William H. Wilson's report that it was also used on East Maui and on Moloka'i, but did not discuss the matter from a historical perspective.

<sup>12</sup>Consistent in one direction at least. However, many instances of /e/ are written with *e* as well.

<sup>13</sup>This is Hervey's explanation of *ai* (1968:26): "In Kauai orthography, <ai> is pronounced [a-i], while in the manner of Pukui and Elbert, *maha* is pronounced [mahea]. Pukui and Elbert indicate that there is a tendency for [a-i] to change to [e-i] in 'fast pronunciations'."

<sup>14</sup>The phrase "just possible" is inexplicable, for the main part of Hervey's dissertation concerns the missionaries' decisions about such consonant alternations.



1777, II:319-322). The published version omitted Cook's introduction, which follows (Admiralty MS 55/108, pp. 243-244, quoted in Lanyon-Orgill 1979:47):

This Vocabulary I had chiefly from M<sup>r</sup>. Anderson Surgeons first Mate, who was indefatigable, in enriching it with all the Words he could collect ... In order to help the reader to a proper pronunciation of the different Words, I desired M<sup>r</sup>. Anderson to draw up such Rules as he thought would answer this end, which he accordingly did ...

In these rules, Anderson took the contemporary approach to orthography: he proceeded from spelling to sound.<sup>15</sup> And he concentrated on the vowels, considering them to be the "regulation of all sounds". Although it is impossible to be sure what Anderson's own pronunciation of English was like (the situation is complicated by the two variables of geographical dialect and time), Table 1 is an attempt to convert his conventions into English phonetics, and from there, into Tahitian phonology.

TABLE 1: A PHONETIC AND PHONEMIC INTERPRETATION OF ANDERSON'S CONVENTIONS

Anderson's transcription	English phonetics	Current TAH orthography
a	[ʌ], [a]	a
<u>ai</u>	[eʏ]	e
e	[e], [ɛ]	e
i	[i]	-i- (?)
<u>ee</u>	[iʏ]	i
y	[aʏ]	ai
o	[oʷ]	o
<u>oo</u>	[uʷ]	u
<u>eu</u>	[yuʷ]	iu
u	[ʌ]	a
ou	[aʷ]	au

Notes: In the chart, an underline represents what was a long stroke or a ligature over the two vowel letters. Anderson's *o* doesn't seem to represent /o/ in today's Received Pronunciation [ɛʷ]. Perhaps it was closer to RP 'ought'.

Anderson discussed the diacritical convention illustrated in the table: a ligature joining two vowels indicated that they were to be pronounced as "one simple sound". Dots over two vowels in succession meant that the sounds they represented were to be "expressed singly" – that is, they constituted two syllables. In his examples, Anderson gave "Ròà" ('great, long, distant'), with a dot over each vowel, probably to distinguish from the English spelling convention of *oa* representing a single sound, as in *boat*. Accent was marked before the syllable in question. Finally, a comma between parts of a word, especially reduplicated portions, represented a "rest or small space of time" ... but not a "full stop".

Table 2 shows Anderson's Tahitian examples (p. 322) respelled according to his conventions.

<sup>15</sup>With the advantage of hindsight, we now look on this approach as quite the opposite of the correct one. But so far as I know, it was not until Peter Duponceau's *English Phonology* (1817) that an attempt was made to reverse this direction and work from sound to symbol.

TABLE 2: TAHITIAN WORDS RESPELLED

Anderson's spelling	Gloss	Respelling	Modern spelling <sup>16</sup>
Ròà	Great, long, distant	<i>roa</i>	<i>roa</i>
E'reema	Five	<i>e'rima</i>	<i>e rima</i>
Ry'poeea	Fog or mist	<i>rai'poia</i>	<i>raipoia</i>
E'hoora	To invert, or turn upside down	<i>e'hura</i>	<i>e huri</i>
Paroo, roo	A partition, division or screen	<i>paruru</i>	<i>paruru</i>

In the Tahitian list, the diacritics were preserved through the printing.<sup>17</sup> However, in the Hawaiian list, which is without any such modifications, one of two things must have happened: either the diacritics were discarded by Anderson himself, or they were used in the manuscript but were lost in the printing process.

Having noticed other instances in which information was lost or confused in the transition from manuscript to printed page, I tend to favour the second solution. The following examples may serve as possible evidence that this is indeed what happened.

Note "Eeeheu" (*ihu*) 'The nose'. Would Anderson have written "eee" as such? It is unlikely. In his directions for pronouncing Tahitian, "e" represented /e/, whereas "ee" represented /i/. Thus, "Eeeheu" would have been fairly close to *he ihu*.

Another example is "Ooma ooma" (*umauma*) 'The breast' versus "Too" (*kō*) for 'Sugar cane'. For Anderson, the single "o" represented /o/, but "oo" represented /u/. In the Tahitian list, one of the transcriptions for 'sugarcane' is "Too", written without italics, perhaps reflecting Anderson's perception of the long vowel.

Still another example is underlined *ai*, which does not represent /ai/, but /e/, as (according to his example) in the second vowel of 'Arabia'. For instance, Anderson wrote "Haire" 'To go'. As it is printed, it looks nearly like Māori *haere*.<sup>18</sup> However, "Haire" would represent *here*, an accurate writing of what is now written as *hele*.<sup>19</sup>

Since the third volume of Anderson's journal of the third voyage is lost (Beaglehole 1967:cx), it is impossible to **prove** that Anderson continued the orthographic conventions described for Tahitian. But the earlier volumes, which cover the period up to 2 September 1777, add weight to the argument, for in them, the conventions are used regularly. For example, Anderson's MS word list from New Zealand, collected in February 1777, shows ligatures over certain vowel combinations, dots over others, and accent marks before the appropriate syllables.

<sup>16</sup>All the Tahitian words were found in Davies 1851.

<sup>17</sup>Although the printer of the account of the second voyage did not do violence to Anderson's system (that is, no information was lost), the original is perhaps more elegant. Any two vowel letters used to represent one sound, such as *ai* for [e], were written with a ligature over them, or – in Anderson's own words – "joined together". This convention was translated by the printer into italics, and the text of the description altered accordingly. Another printer's convention was the substitution of a diaeresis over the first member of a pair of vowels, rather than a dot over each, to indicate that they belonged to separate syllables.

<sup>18</sup>Although Lanyon-Orgill (1979:47-48) quoted Anderson's guide to the pronunciation of Tahitian verbatim, he apparently did not realise that the conventions applied to Hawaiian as well. Thus, he interpreted *Haire* as *haele*, a dual or plural form for 'come'.

<sup>19</sup>Hervey (1968:26) went into an elaborate (and, incidentally, backwards) explanation to relate another *ai* spelling and an /e/ pronunciation.

Understanding Anderson's conventions turns a good many confused-looking forms into words that resemble those in the current spelling. This is not to say that his list is entirely accurate, but it is not a bad showing, considering the circumstances under which it was collected.

Table 3 shows the first twenty-five words from Anderson's Hawaiian list, with a possible respelling according to the Tahitian conventions compared with the modern equivalents.<sup>20</sup>

TABLE 3: SAMPLE FROM ANDERSON'S HAWAIIAN WORD LIST, RESPELLED

	Hawaiian	Gloss	Respelling	Modern spelling
1.	Tehaia	Where	<i>kehea</i>	<i>i hea</i> (specifically)?
2.	Mahaia	ditto	<i>mahea</i>	<i>ma hea</i> (generally)?
3.	Aorre, or Aoe	No	<i>aole</i>	<i>'a'ole, 'a'oe, 'a'ohe</i>
4.	He oho <sup>21</sup>	The hair	<i>he oho</i>	<i>he oho</i>
5.	E poo	The head	<i>e poo</i> <sup>22</sup>	<i>he po'o</i>
6.	Pepaiee aoo	The ear	<i>pepeiau</i>	<i>pepeiao</i>
7.	Heraee	The forehead	<i>herae</i>	<i>he lae</i>
8.	Matta	The eye	<i>maka</i>	<i>maka</i>
9.	Pappareenga	The cheek	<i>papalinga</i>	<i>papālina</i>
10.	Haieea	Fish	<i>heia</i>	<i>he i'a</i>
11.	Eeeheu	The nose	<i>eihyu</i>	<i>he ihu</i>
12.	Oome oome	The beard	<i>umi umi</i>	<i>'umi'umi</i>
13.	Haire	To go	<i>hele</i>	<i>hele</i>
14.	Erawha	Tears of joy	<i>elawa</i>	? <i>le'a</i>
15.	Aee	The neck	<i>ai</i>	<i>'ā'ī</i>
16.	Poheeve	The arm	<i>pohivi</i>	<i>po'ohiwi</i> 'shoulder'
17.	Ooma ooma	The breast	<i>uma uma</i>	<i>umauma</i>
18.	Heoo	The nipple	<i>heu</i>	<i>he ū</i>
19.	Peeto	The navel	<i>piko</i>	<i>piko</i>
20.	Hoohaa	The thigh	<i>huha</i>	<i>'ūhā</i>
21.	He, wawy	The leg	<i>he wawai</i>	<i>he wāwae</i>
22.	Eroui	Wait a little	<i>elauī</i>	—
24.	Myao	Finger and toenails	<i>maiao</i>	<i>mai'ao</i>
25.	Eeno	Bad	<i>ino</i>	<i>'ino</i>

<sup>20</sup>T and rare changed to *k* and *l*, respectively.

<sup>21</sup>Note that the majority of nouns occur with *he*. In a sense, each phrase is a discourse, and new information is introduced as such (Hawkins 1979).

<sup>22</sup>This is an example in which we assume that Anderson made a distinction between "oo" and "oo".

## 6. RE-EXAMINING THE LIST

Rewriting Anderson's word list according to the conventions that were almost certainly lost in the editing and printing gives his work on Hawaiian a credibility that was missing before. Now perhaps the list should be re-examined for what it can tell us about the Hawaiian spoken on Kaua'i in 1778.

### 6.1 PHONOLOGY

It was noted above that Anderson's consistent transcription of *t* shows that that sound was used regularly in at least some parts of Kaua'i, and was not interchangeable with [k], as more recent records show (Newbrand 1951). But this is not particularly new information, since the missionary records contain many references to the predominance of [t] over [k] on Kaua'i. As for the shift to [k], Judd, Pukui and Stokes (1945:13) noted that [t] was more common "between 1778 and 1809 according to the dozen vocabularies made in those years. By the time the present Hawaiian orthography was established in 1825,<sup>23</sup> the "k" sound had become so general that the character "k" was adopted."<sup>24</sup>

Anderson also consistently wrote *r*, as opposed to the *l* that is written today. As for the phonetic nature of the sound, a careful reading of the missionary correspondence in the 1820s and later reveals that it was a tap – rather like the *r* in Spanish, Tahitian or Māori – and not an *r* like that in American English.

Another feature of pronunciation reflected in the list, however, is much less well known: there are five examples of *ng* spellings, apparently representing Anderson's hearing of [ŋ]: *Pappareenga* (*pāpālina*) 'cheek', *Tangaroa* (*Kanaloa*) name of god, *mango* (*manō*), 'shark' – and two examples showing alternation: *tanata* ~ *tangata* (*kanaka*) 'person' and *moena* ~ *moenga* (*moena*) 'mat'.

Treating Anderson's list as a serious document, Geraghty (1983:557) interpreted these data as showing that Proto Polynesian \*ŋ, was usually reflected as [n], but occasionally as [ŋ]. Such information would have added depth to those few treatments of Hawaiian phonology that exist.

Hervey (1968:28) apparently dismissed the idea that Anderson's transcription of *ng* was correct: the only example he treated is *pappareenga* 'the cheek', with the comment "The spelling symbols <ng> could indicate that Anderson thought he had heard [-iŋa] or possibly [-iŋga]." *Ng* does not appear on Hervey's list of "consonant graphemes" (p. 25).

Three unusual transcriptions of *h* – *eehehu* 'the nose', *eahoi* for *he ahi*, 'the fire', and *erahoi* for *he lahi* 'thin' – suggest that the Hawaiian pronunciation of /h/ at the time may have included a palatal constriction common in certain dialects of Māori spoken around the turn of the nineteenth century and in the Marquesas at least as late as the 1930s.<sup>25</sup>

### 6.2 VOCABULARY

Out of the nearly 240 words on Anderson's list, there are still over two dozen not yet identified. Either they are forms that have since dropped out of general Hawaiian, or they are Kaua'i words not

<sup>23</sup>This is a mistake; the correct date is 1826.

<sup>24</sup>This observation, however, dealt only with the variable of chronology, and not with geographical distribution.

recorded elsewhere. Although we have little information about vocabulary differences in Hawai'i at that time,<sup>26</sup> Bennett (1930:59-61) listed a number of artefacts, described but unfortunately unnamed, "distinctively or predominantly found on Kauai".

One botanical term poses an interesting problem. Anderson collected the word *Tearre* (representing either *kiale* or *kiele*), which he identified rather loosely as 'Gardenia, or Cape Jasmine'. Now, however, *kiele* has come to refer to the gardenia that was introduced after European contact. Note the pertinent part of the dictionary entry for *kiele* (Pukui and Elbert 1986, apparently based on Neal 1965):

1. n. Gardenia (*Gardenia augusta*), introduced. (Neal 799-800.) (PPN *tiale*.)
2. vi. To emit fragrance; to perfume with *kiele*, as garments.

The fact that a native gardenia, *kiele* (or *kiale*), existed at the time of European contact seems to have been missed by lexicographers and botanists alike. In Andrews's (1865) dictionary entry for *kiele*, we find a suggestion that the plant was not recently introduced:

The name of an odoriferous shrub or tree; he laau aala. Some say it was brought from a foreign country, but the word is found in two ancient meles at least. [The two *meles* are quoted.]

This information was included in the 1922 edition, but in no later dictionaries.

It is suggested (Derral Herbst, pers. comm.) that perhaps *kiele* was an alternate term for *nānū*, a native species of gardenia, but unrecorded since Anderson. Still, closer attention to Anderson would have cleared up the confusion in the current dictionary: an introduced plant whose name has a Proto Polynesian pedigree.

## 7. CONCLUSION

Ironically, Anderson's word lists from the third voyage seem to have suffered at every touch of an editor and printer. First, Dr John Douglas,<sup>27</sup> who edited the official publication of Cook journals (1784) and used much from Anderson's journals as a supplement (Beaglehole 1967:vi), printed the Hawaiian word list with no diacritics whatsoever, thus changing a fairly accurate rendering to the "defective" vocabulary referred to in section 4.

Next, in his editing of the Cook journals, Beaglehole (1967:817) reproduced (from the original manuscript, one presumes) Anderson's Māori list of 21 words, including dots over vowels and accent marks, but excluding the all-important ligatures over *ai*, *ou*, *ee*, and *oo* to show that each was to be pronounced as a single sound. Beaglehole, although he partially understood Anderson's system,<sup>28</sup> dispensed with this convention for what I consider an inadequate reason: because it would "often have meant a most unsightly page" (1967:ccxviii). In certain forms that would otherwise have

<sup>25</sup>Kendall's Māori grammar (1815), reflecting the Bay of Islands dialect, contains many instances of /h/ written as *sh*. Elbert (1941:58) mentioned a similar sound still used in the Marquesas in the 1930s. Note that in the only examples from Anderson, /h/ is followed by a high vowel.

<sup>26</sup>Kotzebue (1821:305) mentioned some vocabulary differences between O'ahu and Hawai'i that caused communication problems.

<sup>27</sup>"Canon of St Paul's since 1776 and fellow of the Royal Society since 1778 ..." (Joppien and Smith 1987:162).

<sup>28</sup>He wrote that the stroke over two vowels represented *quantity*; actually, Anderson did not generally recognise quantity, but used this convention to indicate *quality*.



been ambiguous (such as a succession of three es), he added the diacritics. But his altered system does not allow for a possible distinction between "oo" (for ō or o'o) and "oo" (for u).

Finally, Lanyon-Orgill (1979:139-140) mentioned three versions of Anderson's short Māori list from the third voyage – the MS, the first printed version (1784), and Beaglehole's transcription of the manuscript – but seems to have chosen the last to reproduce, deleting one more diacritic in the process. Thus, we have a progression from (for example) "Makka'reede" to "Makka'reede" and finally to "Makkareede".

Lamenting the loss of Anderson's final volume, Beaglehole wrote (1967:cxci): "... and we should like our Anderson entire." So should we, especially with respect to the diacritics that change a "defective vocabulary" to a valuable document that gives us new insight into the nature of the Hawaiian language at the time of the first European contact.

## BIBLIOGRAPHY

- ANDREWS, Lorrin, 1865, *A dictionary of the Hawaiian language, to which is appended an English-Hawaiian vocabulary and a chronological table of remarkable events*. Honolulu: Henry M. Whitney. (Revised (1922) by Henry H. Parker. Honolulu: Board of Commissioners of Public Archives of the Territory of Hawaii.)
- BEAGLEHOLE, J[ohn] C[awte], ed., 1967, *The journals of Captain James Cook on his voyages of discovery*, vol. 3: *The voyage of the Resolution...* Cambridge: Hakluyt Society.
- 1974, *The life of Captain James Cook*. [Journals... series, vol. 4.] Stanford: University Press.
- BENNETT, Wendell C., 1930, Kauai archaeology. *Papers of the Hawaiian Historical Society* 17:53-62.
- 1931, *Archaeology of Kauai*. Bernice P. Bishop Museum Bulletin 80. Honolulu: The Museum.
- COOK, James, 1777, *A voyage towards the South Pole and round the world, performed in His Majesty's ships, the Resolution and Adventure, in the years 1772, 1773, 1774, and 1775...* 2 vols. London: W. Strahan and T. Cadell.
- COOK, James and J. KING, 1784, *A voyage to the Pacific Ocean, undertaken by the command of His Majesty, for making discoveries in the Northern Hemisphere...*, volume 3, Appendix V: Vocabulary of the language of Atooi, one of the Sandwich Islands. London: W. and A. Strahan, for G. Nicol & T. Cadell.
- DAVIES, John, 1851, *A Tahitian and English dictionary, with introductory remarks on the Polynesian language, and a short grammar of the Tahitian dialect...* Tahiti: London Missionary Society Press.
- ELBERT, Samuel H., 1941, Chants and love songs of the Marquesas islands, French Oceania. *Journal of the Polynesian Society* 50:53-91.
- FREYCINET, Louis Claude de Saulses de, 1978, *Hawai'i in 1819: a narrative account by Louis Claude de Saulses de Freycinet*. Translated by Ella L. Wiswell, notes and comments by Marion Kelly. Pacific Anthropological Records No. 26. Honolulu: Department of Anthropology, Bernice P. Bishop Museum.
- GERAGHTY, Paul, 1983, Review of Peter A. Lanyon-Orgill *Captain Cook's South Sea Island vocabularies*. *Journal of the Polynesian Society* 92/4:554-559.
- HAWKINS, Emily A., 1979, A new look at the determiner system in Hawaiian. Paper presented at the 5th Annual Linguistic Society of Hawaii Oceanic Festival, 3 May 1979, Honolulu, University of Hawaii.
- HERVEY, Wesley David, 1968, A history of the adaptations of an orthography for the Hawaiian language. Ph.D. dissertation, University of Oregon.
- JOPPIEN, Rüdiger and Bernard SMITH, 1987, *The art of Captain Cook's voyages*, volume 3: *Catalog: The voyage of the Resolution and Discovery 1776-1780*. Melbourne: Oxford University Press.
- JUDD, Henry Pratt, Mary Kawena PUKUI and John F.G. STOKES, 1945, *Introduction to the Hawaiian language (an English-Hawaiian vocabulary, comprising five thousand of the commonest and most useful English words and their equivalents, in modern Hawaiian speech, correctly pronounced, with a complementary Hawaiian-English vocabulary)*. Honolulu: Tongg. (New edn, 1955.)
- KENDALL, Thomas, 1815, *A korao no New Zealand; or, the New Zealander's first book; being an attempt to compose some lessons for the instruction of the natives*. Sydney: G. Howe.



- KOTZEBUE, Otto von**, 1821, *Voyage of discovery in the South Sea and to Behring's Straits, in search of a north-east passage, undertaken in the years 1815, 1816, 1817 and 1818...* Translated by H.E. Lloyd. 2 vols. London: printed for Sir Richard Phillips. (Da Capo Press edition 1967.)
- LANYON-ORGILL, Peter A.**, 1979, *Captain Cook's South Sea Island vocabularies*. London: the author.
- NEAL, Marie C.**, 1965, *In gardens of Hawaii*. Bernice P. Bishop Museum Special Publication 50. Honolulu: Bishop Museum Press. (New and revised edition first published in 1928.)
- NEWBRAND, Helene L.**, 1951, A phonemic analysis in Hawaiian. M.A. thesis, University of Hawai'i.
- PUKUI, Mary Kawena and Samuel H. ELBERT**, 1957, *Hawaiian-English dictionary*. Honolulu: University of Hawaii Press.
- 1986, *Hawaiian dictionary: Hawaiian-English, English-Hawaiian*. Revised and enlarged edition. Honolulu: University of Hawaii Press.
- SMITH, Bernard**, 1984, *European vision and the South Pacific, 1768-1850*. (2nd edn; 1st edn 1960, OUP.) Sydney: Harper & Row.
- WALCH, David B.**, 1967, The historical development of the Hawaiian alphabet. *Journal of the Polynesian Society* 76:353-366.
- WISE, Claude Merton and Wesley HERVEY**, 1952, The evolution of the Hawaiian orthography. *Quarterly Journal of Speech* 38/3:311-325.

MASS AND COUNT NOUNS IN CHINESE AND ENGLISH:  
A FEW FURTHER WHORFIAN CONSIDERATIONS

RON SCOLLON AND SUZIE SCOLLON

1. INTRODUCTION

Arguments, Whorfian or otherwise, if one plays by the rules, are most often developed either deductively from thesis to supporting evidence or inductively from marshalled evidence to the final all but inevitable conclusion. There is a third logic of development, however, which we intend to use in this case. We plan to state our case and then run for cover, leaving as much confusing evidence lying around as necessary to conceal our retreat. It is a hit-and-run logic, a kind of rhetorical commando raid on an all but impregnable but strategically central question: can we learn anything about how people think by studying the characteristics of the language they habitually use? In this case the specific form is this: can we learn anything about the differences between Western ways of thinking and Chinese thought by studying differences between English and Chinese?

We begin with the thesis: nouns in Chinese are mass nouns, in Indo-European nouns are (predominantly) count nouns; this distinction has given rise to Platonism in Western philosophy and to the philosophies of Confucianism and Taoism in China. This idea is not ours but that of Chad Hansen (1983). He advances this idea as a way of accounting for the widely acknowledged but never satisfactorily depicted difference between Chinese and Western thought. We think Hansen is right and that we have good ethnographic evidence to support his position but...

First of all, we should try to clarify what Hansen is saying. Hansen argues that Platonism arose in Greek thought, and continues as the philosophical core of Western thought to this day, because of the problem of set and member grounded in the count nouns of Indo-European languages. The two crucial aspects of count nouns are that they take pluralisation ('a horse', 'some horses'), and that they take some form of article ('a horse', 'the horse'). Given the philosophical problem of determining what is a horse, to take an example not at all at random, to the Platonic mind no particular horse can suitably be the definition of the idea of horse, because no one horse has all and only the characteristics of 'horseness'. Given a set of many members, one only arrives at 'horseness' through abstraction of those qualities that form the defining characteristics, four legs (except in pathological cases) and the rest. Hansen's argument, much fuller and more credible than we are able to demonstrate here, is that because we are constantly troubled to determine whether we are talking about one particular horse or several, we are also, perhaps unconsciously, troubled to know whether we are talking about particular horses or 'horse' in general. This constant push towards making a

distinction between one and many, and then also between particular and general, gave rise in Western thought, with Platonism as the paradigmatic case, to a general concern for distinctions of set and member.

In Hansen's thinking, the concept of set and member is the paradigmatic case in Western thought for linguistic meaning. Certainly the philosophical literature attests to a longstanding precedent of using count nouns as examples in arguments about reference. Platonic idealism is, for Hansen, a response to the intuition that there is something held in common by all members of a set which determines a prototypical or ideal member.

Hansen then contrasts Chinese nouns, which he feels can be best understood by Westerners as being like the mass nouns in Indo-European languages (such as Classical Greek and Modern English), with count nouns. In an extended argument, drawing principally on classical philosophical texts, but not ignoring questions which arise in Modern Chinese,<sup>1</sup> Hansen then argues that nouns in Chinese are mass nouns, not unlike English 'water', 'mud' or 'sugar'. What is significant about a mass noun is that matter is conceived of as an extended 'stuff' (Hansen's wording). The crucial questions which arise, then, are not regarding the ideal characteristics which define the stuff, but the boundaries between one kind of stuff and another kind of stuff. In other words, Hansen believes that Chinese mass nouns have given rise to an ontology of form and substance rather than an ontology of idea or prototype.

Where the noun 'water' is in question, we Westerners have no problem. East or West, water is a mass noun, and we have no trouble with thinking of it as taking the shape of its container, being spoken of in container-sized bits, and as being essentially the same stuff in all its instantiations. A horse or a man is another matter. In English, if we contrast 'a horse' and 'horse', most likely we'd think of the latter as meat. Fish, perhaps, is a better example, where 'a fish' means one animal but 'fish' means the flesh we eat for dinner. In Hansen's view, in Chinese 'horse' and 'man' are like 'water'; they are extended 'stuffs' or substances which take the shape of particular containers. This is an intriguing view, and once one gets around to trying to think that way, it seems to make a different kind of sense of much of Chinese philosophical thought and life. That is what attracts us to this idea.

Not least among Hansen's arguments is that this conception puts Chinese philosophers and the Chinese language on a somewhat better footing than they have sometimes been allowed by Western thinkers in the past. In Hansen's view it is not fair to say that Chinese philosophers were not able to conceive of the set-member distinction nor ideal types, but just that they had little interest in doing so. On the other hand, Platonic thinkers had little interest in typical Chinese philosophical issues and concentrated on questions of definition, reference and truth. In the past many Westerners have troubled over the absence in Chinese of both a distinction between concrete and abstract and the distinction between singular and plural, two distinctions which do not seem to bother us in reference to Indo-European mass nouns – it apparently only bothers us where Chinese have a mass noun in places where we have a count noun. Looked at from this perspective, if we genuinely accepted the world view implied by a total mass-noun inventory, it might ultimately be more difficult to

---

<sup>1</sup>As an example of what sort of problem we mean, in contemporary Mandarin some nouns which are mass nouns in English such as 'water' may be used without the individuator or measure word. On the surface of it this may indicate a distinction between mass and count categories not unlike that found in Indo-European. We take up this question in more detail below.

understand how we can consider two horses to be different things since they are both, like two cups of water, just two skins full of horse.

Hansen's notion is that not only are all Chinese nouns mass nouns, but that Chinese philosophy is fundamentally nominal; "behavioral nominalism" is what he calls it. One can use the characteristics of mass nouns as a kind of philosophical template by which to understand Chinese thinking, and ultimately, we presume, the sort of ethnographic matter of which we have been trying to make sense. If this were an argument, rather than a sketch of one, a considerable amount of detailed evidence would be necessary at this point to draw a clear picture of that philosophical template. In place of that evidence, the following table of contrasting points can suggest the outline:

Mass Nouns	Count Nouns
Definition by example	Definition by analysis
Discrimination	Description
Boundary marking	Prototype
Interpretation	Observation
Taoism	Nihilism

## 2. DEFINITION BY EXAMPLE, DEFINITION BY ANALYSIS

The count-noun ontology gives rise to a semantics in which knowing a word seems to centre on definition by analysis. Socrates, for example, dismisses his opponents whom he says do not know what a word means since they are unable to give an unequivocal analysis of its meaning.<sup>2</sup> Language is thought of as describing a world independent of language; meanings of words seem to cluster about a prototypical example, or in the extreme Platonic form, an ideal type. This leads to a general sense that reality is an idea or a prototype from which instantiations depart. Along with this is a general thrust towards the making of statements, or descriptions of the world, which are judged for truth. Truth itself tends to be established by analytical rules, and forms the goal of philosophical activity.

The semantics of mass nouns centres on discrimination and prefers definition by example. As a result, the crucial questions depend on boundary marking. In that sense, the ontology of count nouns seems to focus towards the centre, the prototype, while the ontology of mass nouns seems to focus towards the boundaries, the forms, within which the substance of reality is contained. Since an undifferentiated substance might be found to be held within an abundance of containers, reality comes to be seen as conventional. What is held in one container might as well be held in another; what is called by one name might in other circumstances be called by another; the substance itself is thought of as unchanged in this process of renaming.

---

<sup>2</sup>One wonders in this sort of argument where those opponents of Socrates came from. By Hansen's argument they are behaving in a very Chinese way for Greeks. Perhaps it is because Socrates has won these arguments that we take his view of language as the prototype, not that of his equally Greek-speaking opponents.

### 3. DISCRIMINATION, DESCRIPTION: THE RECTIFICATION OF NAMES

Since in the logic of mass nouns the discrimination of the names of things is seen as conventional, the choice of which name one uses comes to be tested against the pragmatic consequences of choosing that name. This pragmatic emphasis in Confucian writings on semantic discrimination is most often translated into English as 'the rectification of names'. It is thought by some to form the essential core of the Confucian philosophy of language (Hall and Ames 1987). The most well-known example occurs in the *Mencius*. The word in question is 'king'. In Confucian thought, one is called the king only when one behaves as a king within the canons of the Confucian code. The killing of a king is prohibited within the Confucian code, and yet the last king in a dynasty is sometimes allowed to be killed. Mencius was questioned about this inconsistency and answered as follows:

King Hsuan of Ch'i asked, "Is it true that T'ang banished Chieh and King Wu marched against Tchou?"

"It is so recorded", answered Mencius.

"Is regicide permissible?"

"A man who mutilates benevolence is a mutilator, while one who cripples rightness is a crippler. He who is both a mutilator and a crippler is an 'outcast'. I have indeed heard of the punishment of the 'outcast Tchou', but I have not heard of any regicide".

(Lau 1970:68)

If a man is called 'king,' he cannot be killed. If a man is called 'outcast,' it is not a king one is killing. The moral consequences follow from the name given. If language can be divided into two very broad functions, the descriptive and the regulative, the semantics of the count noun comes down on the descriptive side and the semantics of the mass noun comes down on the regulative. If the count-noun logic drives in the direction of the establishment of true statements, the mass-noun logic drives in the direction of examining the social and moral consequences of one's statements.<sup>3</sup>

### 4. BOUNDARY MARKING, PROTOTYPE: LOCAL PATTERNING

The rectification of names is the pragmatic outcome of the Chinese mass-noun philosophy. For our purposes, however, the most interesting aspect of these contrastive mass-noun and count-noun philosophical templates is the absence in the mass-noun logic of any interest in definition by logical attributes or prototype and the contrasting great concern with discriminations and boundary marking. Hansen notes that:

nowhere in any of the traditional accounts of the rectification of names was there any mention of definitions, ideals, ideas, ideal types, concepts, essences, or even of meanings. (1983:82)

What is crucial in this mass-noun logic is the discrimination of boundaries. In nouns, as in geopolitics, the Great Wall is the symbol of China.

<sup>3</sup>This aspect of the rectification of names is all too familiar in the modern context. We find that no students were shot, only ruffians and radicals, or even vermin. Unfortunately, this practice of semantic readjustment is not the exclusive property of speakers of mass-noun-only languages. Our point, and Hansen's, is only that English (or Greek) and Chinese differ in their relative emphasis on the descriptive and regulative functions of language. A mono-functional utterance in any language would be a difficult thing to establish indeed.



Western images of China abound. Two which come to mind immediately seem at first to form an irreconcilable contrast. A Tang Dynasty landscape painting in which clouds hide all but glimpses of mountains is one image held in Western imaginations, while another is that of a highly ornate brocade, or perhaps a temple. The contrast is between an aesthetics of the barest suggestion and one of an extremely dense symbolic complexity. On the more mundane level one sees in the streets of Taipei, Hong Kong, or even Zhongshan a person dressed in slacks or a skirt of one striking and complex pattern and a shirt or blouse of a different (and, to Western eyes, glaringly clashing) complex pattern. In contrast to a Western, or even Japanese, aesthetics of sparsity of composition and form, one sees in China patterns of juxtaposed intricacy which strike a Western observer as lacking in all sense of balance and perspective.

And perspective may be, indeed, the point. Clothing patterns, temple designs, the furnishing of modern houses, Tang Dynasty landscape paintings, and a nine-course Chinese dinner all reflect what we think of as local patterning. The aesthetic organisation which applies within a boundary is not carried over to apply across the boundary. In a Western analysis this might be called an absence of perspective. The landscape painting does not resolve into one, organising perspective. There are local areas of patterning and perspective which carry no implications for patterns in other areas. The size of the pavilion perched high on a ledge bears no particular relationship to the size of the boat tied up to a bank in the lower corner of the same painting. The design and colour relationships of the material from which the pants are made bears no relationship to those of the shirt. Each answers to the logic of a local aesthetic pattern.

In this, as in so many other things Chinese, the written character seems to form a paradigmatic example. With extremely rare exceptions,<sup>4</sup> the Chinese character is calligraphed within an imaginary square boundary. Within that boundary there is a time-honoured and rather complex aesthetic of strokes, both position and order, which applies to all characters whether they are written with one stroke or with sixty-four. There is no Chinese equivalent to the Western alphabetic practice of proportional spacing, or kerning, of letters such that an 'i' is given less space than an 'm', or a 'w' is given more than other letters.

The gist of the idea of local patterning, or internal patterning, is that the boundary forms an aesthetic or conceptual limit to the concern for order or harmony. A row of characters has no flow from one to the next;<sup>5</sup> in a painting there is no single perspective; food is prepared and served in small morsels; if a string of elements is connected one to one to one, there is no implication that the first and the last must also relate to each other. In a negative (Western) assessment there is no sense of perspective nor any dominating point of view. A linked chain is the dominating metaphor. The cultivated practice of linking couplets of poetry in a gathering of poets is a further instance, we think, of local patterning. The point in these poems is not to create a grand work of art, with economy of expression and a central theme, but more on the model of ping-pong, to catch the other poet's phrase, match it in mood, imagery and rhyme, and then to pass it on to the next poet in line. The aesthetic pleasure derives from a skilled control of the line or the couplet, not from the creation of a unified

---

<sup>4</sup>In the National Palace Museum in Taipei we were able to locate only one example of calligraphy in which characters were arranged with an overflow of one character into those both above and below it and those on either side.

<sup>5</sup>We are aware of 'grass' writing, of course, in which characters are written in a highly flowing style. Nevertheless, the examples in which any one grass character flows over into another are still quite rare. This is not the case so far as we know in grass style as practiced in Japanese calligraphy.

A style of calligraphy developed in the 20th century in which calligraphed words appear superimposed upon wallpaper-like floral patterns corresponds quite closely with the use of patterns in Chinese styles of Western clothing.



perspective. The Western film technique in which a montage of unrelated elements is juxtaposed was introduced by Eisenstein on the inspiration of juxtaposed images in Chinese poetry (Yip 1976).

In designing a new campus for Providence University in Shalu, Taiwan, each department was given a separate building with markedly differing architectures. Although a single architect was employed to create these pockets of local patterns, he was instructed to follow the tastes and inclinations of the deans and department heads in creating this aggregate of designs. This juxtapositioning of unrelated architectural elements contrasts strongly with Tunghai University just a short distance away which, being designed on a Japanese model, follows continuous motifs throughout the campus.

Local patterning and the linked chain of patterns also shows up in the Chinese essay. Moving from one topic to another topic means showing the boundary between topics. In contrast to this, in Western essayist discourse moving from topic to topic requires the re-establishment of the central or dominating perspective. We expect an essay or an argument to centre on a theme: 'Thesis X, Y, Z has three main elements, X, Y and Z'. Then as we begin our development we want to say, 'The first point in support of XYZ is X'. As we make our first transition we would like to see, 'Not only does X support XYZ but Y does as well'. Then in conclusion we want, 'Moreover Z provides the most decisive support to XYZ'. Each shift from X to Y to Z re-establishes the centrality of XYZ and shows where each element fits, much as in a Western painting we expect each element of the painting to point to the vanishing point.<sup>6</sup>

A Chinese essay proceeds from point to point, beginning, of course, with established precedent and ending with a suggestion of the author's point. 'X has said Y. This is interesting and worthy of our attention. Which leads me to think of P. P, of course, makes one think of A. One can never bring up the topic of A without being led immediately to think of M. However, the author believes that M might be modified to say Y+Z'. Each element in the argument, X, Y, P, A and M is treated independently with appropriate quotations from the classics and even personal anecdotes. What one does not do is cast a unifying perspective over the whole and demean the reader by showing the connections among the elements mentioned.

A consequence of local patterning in Chinese discourse is the topic umbrella. Once a topic is mentioned, it is no longer given explicit statement within that unit of discourse. The boundary is made and then discourse continues within the boundary until further notice. As a result, Chinese discourse tends to consist of a series of chain-linked predicates.

## 5. INTERPRETATION, OBSERVATION: INSIDE/OUTSIDE

A major consequence of boundary marking and local patterning in Chinese thought is that there is a relatively higher emphasis on interpretation than on observation. This can also be phrased as a relative lack of response to context in Chinese thought. By way of contrast one thinks of examples beginning with the Parthenon and the Socratic dialogues and coming down to the present laboratory studies in science as an attempt to organise reality from the point of view of a clear, unobstructed perspective. One clears the ground from around a work of architecture so that the eye is not cluttered

---

<sup>6</sup>Of course, vanishing point perspective in Western painting was not used until the Renaissance and has all but disappeared, except perhaps in greeting cards, in the modern period. But to use the painting of the period from the Renaissance to the modern as a prototype is an example of how count-noun thinkers work, so we have decided to let this stand as an example if not as an argument.

up by competing forms; one isolates a phenomenon in the laboratory so that only those aspects one wishes to study come into the picture.

This Western clearing of the ground could be seen as an attempt to control the embeddedness or context of things and events, and so might appear to be not unlike Chinese boundary marking, which attempts to make a clear distinction between what is inside (and therefore subject to local patterning) and what is outside (and therefore ignored). The difference lies, we think, in that in the West, the focus is on the central phenomenon, as a sort of logical vanishing point, and the context is pushed out of the way as much as possible, or until no longer convenient. It is not the boundary which really matters, so much as just getting the peripheral vision cleared of competing perspectives. In the Chinese view the focus is on the boundary. Discriminating a boundary is not only a logical or a descriptive activity, it is a regulative and moral activity. What is outside a boundary is not relevant in any way to what is inside. It is safely (even compulsorily) ignored.

One finds in colleges in Taiwan, for example, a sense of unreality as if the world outside the walls were of no relevance to the activities being pursued within the walls. The entrance exam, of course, sets the entry boundary for Chinese students. American teachers in Taiwan are struck by the extreme emphasis placed on these exams for college entrance as well as for successful graduation. If it were just the exams, however, one might think of parallels with the British academic model rather than with American colleges. Many other activities, however, cluster around these crucial boundaries. The senior year seems more taken with photographs, preparation of memorial writings, and graduation preparations than with the studies that one would think were the prerequisite for graduation. At the same time, and in contrast to American college students at least, it seems little interest is paid to the intellectual aspects of being a student. There is a perfunctoriness in studying that gives the appearance more of role-playing than of intellectual inquiry. It is widely advertised in Taiwan, for instance, that once one is accepted by a major university everything is easy. No study is required to graduate and, because one has graduated from a major university, one is virtually guaranteed a successful life. That one's activities in college should bear any connection to one's work later outside college seems to rarely be considered.

Ignoring context is, of course, an act of interpretation. That might help to explain the Chinese capacity to simply not see what is not interpretable or not relevant to the current local pattern. A man scrupulously goes down the boundary of his tiny yard and the outside of his house. He picks up trash as he goes and tosses it a few feet away onto the property of another person. It is no longer trash cluttering up his space, and so it is no longer a problem. A temple is built so as to open out onto the vista of a stream. The stream becomes filled with the flotsam and jetsam of a rapidly developing country specialising in plastics. A temple looking out onto a stream is an ancient image of tranquillity and beauty. Does a Chinese who has accepted that symbolism actually see the trash? We would say that in a Western observational sense he or she must be able to see it. But the tranquil temple image does not allow the interpretation of ugliness. So a photograph of the garbage-glutted stream and temple taken by a foreigner only elicits the response, "What is that?" The question is not a literal, observational question. The meaning is, "What motives do you have in taking such a picture?" In other words, the questioner wants to know what interpretation one is driving at. The drive to interpretation and its consequences either undermines or motivates all observation. It appears that once the interpretation is made, what falls outside the boundaries is simply not seen. The Chinese emperor who, upon seeing such non-Chinese animals as giraffes and who ordered that the boats of exploration which had gone out to Africa be burned, is an example of this Chinese disinterest in what lies outside one's boundaries.

## 6. INSIDE/OUTSIDE

This distinction between inside and outside governs interpretations in virtually every domain, from calligraphy and clothing styles through kinship relations and the building of houses, to the building of the Great Wall. The inside relations, for instance, are those established when one is born, such as the relations with other members of one's family; they also include those established by being members of the same class in school, entering a company as part of the same groups of recruits, or perhaps serving in the military together.<sup>7</sup> Outside relationships are those strictly utilitarian relationships one has with taxi drivers, ticket collectors, bank tellers, or business clerks. In discourse this distinction governs such matters as who speaks first and who introduces the topics of a conversation.<sup>8</sup> This is not to say that the distinctions between inside and outside are all that obvious to one who does not know the system. The wife in the Confucian system is inside (literally 'inside person') but the wife's relatives are the outside relatives. As in most cultural matters the distinction between inside and outside is what is patterned throughout the culture, not the particular contents of the pattern.

## 7. TAOISM, NIHILISM: SEMANTIC BREAKDOWN

Finally, one way to look at a cultural pattern is to see how it breaks down. As a kind of perceptual archaeology, sometimes in dissolution one sees the former pattern more clearly. When we compare the mass-noun philosophical template with the count-noun philosophical template the question we want to ask is: What happens when the ability (or the desire) to discriminate is lost (or abandoned)? Also, what happens when the ability (or the desire) to achieve clear, definitive prototypical sets breaks down? In other words, what is the shape of mass-noun scepticism as compared with count-noun scepticism? As a quick answer, we suggest that when one loses the discrimination of forms, one is left with the substance; when one loses the definition of the ideal characteristics required for set membership, one gets an undifferentiated world of individuals without any principled groupings. The first case, the mass-noun case, leads to something like Taoism; the second case, the count-noun case, leads to something like nihilism.<sup>9</sup> In other words, we think if someone were to take the time to work it out, it would be possible to show that even though Taoism and nihilism appear to share a great deal in their generalised relativism, Taoism and nihilism differ radically in their psychological valency. A substance without form is still, after all, something, a positive. An individual item in the universe without set membership is, in fact, nothing.

Whorfian arguments are dicey no matter how nicely ethnographic and philosophical speculations line up behind them in support. In the case of Hansen's very Whorfian notion that so much of the Chinese philosophical template has derived from the mass-noun character of the Chinese language, one is reminded immediately of Whorf himself. In drawing a comparison between Hopi and English, Whorf had occasion to use mass nouns in English to draw what, on the surface of it, appears to be quite a contrasting conclusion.

---

<sup>7</sup>These relationships established in schools, work and the military are sometimes talked about as *tong* relationships.

<sup>8</sup>Two systems of conversational initiations and topic control operate in which outside relationships are strictly utilitarian but inside relationships are dominated by the Confucian benevolence-respect pattern. In such cases as between teacher and student, parent and child, or employer and employee the right to initiate conversations and to introduce the topics is reserved for the person in the upper or benevolent position. For a further analysis of this discourse implication of the inside-outside distinction in Chinese thought see Scollon and Wong-Scollon (1990).

<sup>9</sup>One wonders whether it is a reflection of this distinction between Taoist nothing and nihilist nothing that we conventionally capitalise Taoism but not nihilism.

In 'The relation of habitual thought and behavior to language' Whorf's argument (1956:134, but written in 1939) is that mass nouns in English give evidence of an underlying (Whorf calls it 'covert') analysis of all matter into its substance and its form, ultimately giving rise to the transcendental characteristics of Western thought. Hansen, of course, has derived that same Platonic transcendental analysis from English count nouns. Whorf argues that the English use of individuator (such as 'cup' in 'cup of water') indicates the fundamental concern in the habitual thought of speakers of SAE (Standard Average European) languages to objectify abstract substance with a form.

Certainly one cannot have it both ways: if there is such a thing as a logic of mass nouns, that logic ought to apply in about the same way whenever mass nouns are used, whether it is in SAE or in Chinese. Surely mass nouns are not likely to be the clearest indicator of Platonic idealism in SAE while at the same time giving rise to a radically non-transcendental philosophy in China. More directly to the point, perhaps, is that when I stir my tea with a spoon it seems far fetched to think that the spoon, as a count noun, is stirring up philosophical Platonism in my mind at the very moment the tea, as a mass noun, is brewing Taoist contemplations.

One can imagine trying to save the story by arguing that it is, in fact, the count nouns which give rise to Platonism in SAE and that the Platonism carries over into our analysis of mass nouns as well. In any event, it is the individuator that forms the crucial evidence for Whorf, and it is just there that one has problems in the Chinese analysis.

The language of Confucius, unlike modern Chinese, is a language of mass nouns without individualizers (or measure words--also called classifiers by some scholars). At that time, according to Norman (1988), individualizers were markedly rare.<sup>10</sup> If that were all there were to it, we could say that the presence of individualizers in English and their rarity in Confucian Chinese shows the pressure of the count-noun template on English (and SAE) mass nouns. Unfortunately, Modern Chinese has all-too-abundant individualizers, more by a large amount than English. Norman argues that they had gone from being rare at the time of Confucius to becoming commonplace by the Nanbeichao Period (AD 420-589). This gives rise to the question of whether or not this establishment of measures in Chinese indicates a rise of the count noun as a covert category in the millenium between Confucius and the Nanbeichao Period. While it might somehow be argued that there was an influence from Indo-European by way of Sanskrit Buddhist texts, this influence beginning in China around AD 480,<sup>11</sup> the sociolinguistic circumstances required for Sanskrit to exert such an influence seem improbable enough to make this explanation rather far fetched.

---

<sup>10</sup>There remains the problem of knowing whether or not the written language of Confucius represents the spoken language of 'the Chinese' in the 5th century BC. These are actually two problems: we do not really know the relationship between written and spoken Chinese for this period, and the texts referred to as the writings of Confucius were written not by Confucius, but by others at some later date. We hesitate to bring up the fact that China at that time, as now, was sociolinguistically rather complex, so much so that generalisations from those few texts to the thoughts of a culture seem almost hopeless. Some measure of rescue from this position seems to be had in the fact that if we can gain any ground in our understanding of the Confucian language it can be asserted that over the centuries this language has exerted a conceptual pressure on Chinese culture that is, perhaps, unmatched elsewhere in human culture.

<sup>11</sup>This influence comes a little late in the millenium to have done much in any event. Bodhidharma, who is said to have brought Buddhism to China, was thought to have been a Persian monk (Wu 1975). His arrival began a period of extensive study of Sanskrit texts and the translation of these texts into Chinese. There are, of course, many legends indicating a much earlier introduction of Buddhism to China. In one case it is said that Confucius himself knew of the teachings of Buddha (Ch'en 1964).

The evidence from language learning seems to argue against any covert category, or psychological reality, for count nouns in Chinese in contemporary times. It is a truism among language teachers that speakers of Chinese are bedevilled in their learning of Indo-European languages by a pervasive inability to get grammatical number straight. One would expect number to be somewhat more learnable in a new language if one had number in one's own language (either overtly or covertly). A second somewhat reciprocal sort of evidence that individuator seem to have nothing to do with a count-noun analysis is that speakers of English (and other SAE languages) appear to have an equivalent difficulty in learning individualizers in Chinese.

We can only conclude that Whorf's introduction of mass nouns in his comparison of English and Hopi is convincing only if we ignore the evidence Hansen calls up from Chinese. Unfortunately, Hansen is also more convincing if we do not look too closely at a comparison with Whorf's examples. That the two arguments are incompatible seems to us strong enough evidence that neither is on the right track.

In a recent reanalysis of the thought and language of Confucius, David Hall and Roger Ames find much to admire in Hansen's general statements about Ancient Chinese language and logic (Hall and Ames 1987). They take exception, however, to Hansen's assertion that the mass-noun character of Chinese lies at the foundation of Chinese thought. While they draw a parallel set of conclusions about that thought, they attribute these characteristics to the Chinese emphasis on the pragmatic role of language in communication over the referential role. Above we argued, following Hansen, that the emphasis on discriminating boundaries in the Chinese mass-noun ontology leads to an emphasis on the regulative functioning of language. Hall and Ames stand this argument on its head, arguing that it is because, from earliest times, Chinese have regarded the pragmatic role of language to be fundamental that they have developed a philosophy of language concerned with discrimination, setting boundaries, and the rectification of names. The fault Hall and Ames find with Hansen's argument is that contrasting the mass-noun philosophical template of Chinese with the count-noun philosophical template of English still implies for both a referential theory of language.

Without citing Hymes (1966), Hall and Ames have introduced a second type of linguistic relativity to the Whorfian argument. In their view the relationship between Chinese thought and language derives functionally rather than referentially. The crucial difference between Chinese and SAE languages, in their view, lies not in the nature of nouns in these languages but in the relative emphasis given to the pragmatic and referential functions of language. It is this difference which has given rise to the philosophical differences Hansen notes, and it is this difference which underlies the ethnographic distinctions we have sought to understand.

The question with which we began was: can we learn anything about the differences between Western ways of thinking and Chinese thought by studying differences between English and Chinese? Now we are in the position to say that while such a contrast has many attractive elements, we believe the attractions of Whorfian thinking lie closer to art than to science. If science and art can be distinguished not so much by their subject matter as by their methods, perhaps we can say that science is driven by its questions towards some resolution in proof, but art is driven by its observations of the world towards the posing of attractive questions. Like in much art, the proof of a Whorfian argument lies in the attractiveness of its questions.



## BIBLIOGRAPHY

- CH'EN, Kenneth K.S., 1964, *Buddhism in China: a historical survey*. Princeton: Princeton University Press.
- HALL, David L. and Roger T. AMES, 1987, *Thinking through Confucius*. Albany: State University of New York Press.
- HANSEN, Chad, 1983, *Language and logic in ancient China*. Ann Arbor: The University of Michigan Press.
- HYMES, Dell, 1966, Two types of linguistic relativity. In William Bright, ed. *Sociolinguistics*, 114-167. The Hague: Mouton.
- LAU, D.C., 1970, *Mencius*. New York: Penguin Books.
- NORMAN, Jerry, 1988, *Chinese*. New York, Cambridge: Cambridge University Press.
- SCOLLON, Ron and Suzanne WONG-SCOLLON, 1990, Some cultural aspects of teaching English to Asian adults. MS, Sogang Institute for English as an International Language, Seoul, Korea.
- WHORF, Benjamin Lee, 1956, *Language, thought, and reality: selected writings of Benjamin Lee Whorf*. ed. John B. Carroll. Cambridge, Mass.: The M.I.T. Press.
- WU, John C.H., 1975, *The Golden Age of Zen*. Taipei: Hwa Kang Bookstore.
- YIP, Wai-lim, 1976, *Chinese poetry: major modes and genres*. Berkeley: University of California Press.





## GRAMMAR, PERCEPTION AND REALITY

STANLEY STAROSTA

### 1. DIFFERING VIEWS OF REALITY

George Grace's 1987 work, *The linguistic construction of reality* (LCR), contrasts two views of language, the 'mapping' view, which he takes to underlie modern linguistics, and the 'reality construction' view, which he proposes as a more realistic alternative.<sup>1</sup> In the mapping view, the categories of language map directly onto external objective reality, while in his "...*reality-construction* view, the imperfectness of our access to knowledge of the real world assumes central importance. Emphasis is placed upon the fact that we do not have direct access to the real world itself, but only to the data about it provided by our senses" (Grace 1987:6). Thus the reality onto which the categories of language map is a subjective one, an internal psychological construct inextricably bound up with and constituted by these same categories of language: "...the representations provided by our languages are regarded as nothing more than models which serve as surrogates for the real world" (Grace 1987:7); "...language may be said to provide a medium in which humans construct realities for themselves, and...it is in terms of these constructed realities that they live their lives and conduct their affairs" (Grace 1987:108).

Unfortunately, the author doesn't provide us with any of the detailed real-language evidence and argumentation which would be required to make his 'reality-construction' view of language convincing. In this paper, I would like to propose an alternative view, one which seems to lie somewhere between the 'mapping' view and Grace's reality-construction approach, and to outline some evidence from real languages that supports this intermediate way of looking at the relation between language and reality.

#### 1.1 GRACE'S VIEW OF THE RELATION BETWEEN LANGUAGE AND REALITY

##### 1.1.1 LANGUAGES ENCODE A PERSPECTIVE OF REALITY RATHER THAN REALITY AS SUCH

In Grace's view, as stated above, languages encode a perspective of reality rather than reality as such. Here are some further quotes which develop this point:

---

<sup>1</sup>This paper was written during my sabbatical leave under fellowships from the Alexander von Humboldt Foundation (first semester) and the Language and Research Fellowship Program (Committee on Scientific and Scholarly Cooperation with the U.S., Academia Sinica, the Interuniversity Program for Chinese Language Studies in Taipei, and the Luce Foundation, second semester). They deserve the credit, and I deserve the blame.

Robert Blust, ed. *Currents in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, 477-500.

*Pacific Linguistics*, C-117, 1991.

© Stanley Starosta

In short, the unity, and therefore the event as an event, is in the eye of the beholder; what constitutes an event is whatever a speaker is able to interpret as one. (Grace 1987:32)

A linguistic expression of itself does not say anything, it specifies a condition of instantiation for that event, and it provides clues as to what context that instantiation of that event is to be placed in, but it does not say anything. (Grace 1987:38)

### 1.1.2 LANGUAGES DIFFER IN THE WAYS IN WHICH THEY CONSTRUCT REALITY

From LCR, it is possible to extract some more specific properties of the linguistically constructed reality Grace posits:

(a) Languages differ in the ways in which they construct reality:

Of the three components of the content, it is the conceptual event which....shows the greatest variation from language to language....And it is, of course, this component in which the linguistic construction of reality is manifested. (Grace 1987:39)

However, as was pointed out, different individuals or different communities of individuals usually have different linguistic repertoires, and it is just as proper to speak of conceptual worlds which are reflected by particular linguistic repertoires rather than by languages. (Grace 1987:108)

(b) The ways in which languages evolve in their encoding of reality is (partially? wholly?) determined by the external environment in which they develop.

(c) The categories in terms of which languages encode reality can differ in arbitrary ways.

I found no such statement as (c) in LCR, but I also found no statement suggesting that there were any genetically determined constraints on the direction in which a language could develop, so it seems to be a reasonable inference. One gets the strong impression from reading LCR that the way that languages construct reality is quite arbitrary and is a completely accidental consequence of the exigencies of the collective experience of the present and previous speakers of the language.

## 1.2 THE LEXICASE VIEW OF LANGUAGE AND REALITY

The views of reality-encoding which I would like to compare and contrast with Grace's views are those assumed by the lexicase version of lexicalist dependency grammar (cf. Starosta 1988), and are referred to in lexicase as 'perspective'. These views are quite close to Grace's in some respects, but closer to orthodox Chomskyan views in others.

### 1.2.1 LANGUAGES DIFFER IN THE WAY IN WHICH THEY ENCODE REALITY

Yes, languages do differ to some extent at least in the way in which they encode reality, and the differences may correlate with the environments in which the languages are spoken. The ways in which languages evolve are to some extent determined by the environments in which their speakers live. In the area of lexical meanings, for example, this position is hardly controvertible especially with reference to the description of new and unique cultural innovations and the creation of modern scientific and technological terminology. There is no way in which human genetics could have anticipated the development of, say, quasars and black holes. Rather, every language has the ability

to add to its vocabulary in unanticipated and unrestricted directions,<sup>2</sup> with new word meanings I think defined differentially and understood prototypically.

### 1.2.2 THERE ARE UNIVERSALS

The differences being conceded, however, I think it must also be conceded that there are significant aspects of grammar and of the kinds of 'perspectives' that a verb can impose on an external situation which do not differ arbitrarily from language to language or language family to language family, but which rather match each other in ways which are too close and peculiar to be attributed to chance, and too arbitrary and unmotivated to be regarded as the result of environmental pressures. The details of this correlation are so specific and systematic that it seems impossible to explain them in any non-circular way as the result of shared experiences; and languages are capable of such rapid and radical changes that it is impossible to maintain that these cross-linguistic similarities are retentions from some proto-human-language ancestor which somehow were preserved when everything else was changing beyond all recognition. It is hard to avoid the conclusion that they constitute a set of genetically determined<sup>3</sup> universal categories.

The recognition of a set of universal linguistic categories is of course one of the cornerstones of generative grammar, a linguistic tradition which Grace finds to be grounded in the erroneous direct mapping view of linguistic reality-encoding, and this recognition constitutes the primary respect in which the lexicase view of reality-mapping differs from Grace's 'reality-construction' hypothesis. Yes, language does encode perception rather than situations, as will be demonstrated below, but no, these ways of perceiving are not totally a product of the cumulative experiences of the users and propagators of the language. True, some of the examples of linguistic universals that have been given within Chomskyan linguistics are heavily theory-dependent, but others which I will outline below are more clearly intersubjectively verifiable, and too concrete to be easily dismissed as the consequences of working within distinct linguistic paradigms.

In the main body of this paper, I will attempt to illustrate the existence of some linguistic universals in the area of verb subcategorisation, and to demonstrate that they are in fact universals of perception rather than of external situations. To do this, I will first define and illustrate some of the lexicase concepts that differ from conventional modern linguistic ideas. This discussion will place special emphasis on the fact that lexicase views language as encoding human perception (a view akin to, though not identical with, Grace's 'reality-construction' view), rather than as encoding external reality directly (Grace's 'mapping'). In the remainder of the paper, I will outline evidence for this view by showing that it allows the capture of more and better linguistically significant generalisations than frameworks such as Chomskyan grammar and Fillmorean case grammar which assume the mapping view.

<sup>2</sup>This ignores the possibility that there are some things that are just totally beyond our ken, so utterly alien to human experience that we really are genetically incapable of grasping them. I am not going to deny this possibility here, but since there is not much I can do about it, I will ignore it henceforth.

<sup>3</sup>Directly or indirectly. Logically, either such categories are 'hard-wired' in the human brain as linguistic categories (either present from birth or preprogrammed to develop later during the process of maturation), or the human system of information-processing is hard-wired in such a way as to make the development of such linguistic categories inevitable. This latter view seems to have gained popularity recently, but it is hard to see how it could be demonstrated in any non-circular way, so I will assume the former view.

## 2. SOME BASIC PREMISES OF LEXICASE

## 2.1 THE TRIUNE SIGN

In Ferdinand de Saussure's original formulation, a language is a system of *signs*, units made up of sounds associated with meaning. Because the sound-meaning association is arbitrary, there is no way to predict one from the other, and so these associations must be memorised. From the lexicase viewpoint, however, the sign must have three parts, not two: in addition to a pronunciation and a meaning, a word has a particular syntactic distribution (minimally a 'part of speech'), and this distribution too must be memorised, since it is not reliably predictable from the sound, the meaning, or the combination of the two. For example, there is no non-circular way to predict from the meanings or pronunciations of *devour*, *dine* and *eat* that one of them will always occur with an object, one never will, and one can swing either way. If two signs differ in either pronunciation, meaning or distribution, then they are separate entities, independent of each other and memorised separately.<sup>4</sup>

A corollary of this view is that when a given root occurs in more than one grammatically significant environment, it is to be analysed as two separate lexical entries.<sup>5</sup> It is the thesis of this paper that such differences in distribution usually if not always correlate with differences in meaning. Since the root is the same, the lexical entries sharing this root typically refer to the same basic external situation, as defined in terms of truth value, but the meaning difference is typically not one of reference to distinct ranges of external situations (the 'mapping' view), but rather of differences in 'perspective', that is, in the ways in which external situations are perceived. That is, the distinct lexical entries encode not one or more external situations directly, but rather one or more specific 'perspectives' of some external situation. This difference in meaning is not characterisable in terms of truth values, since the two lexical items may well refer to objectively identical external situations but to differing psychological constructs, and this conclusion fits together with Grace's idea that language does not map directly onto a physical reality, but rather reflects a psychological reality.<sup>6</sup>

Obviously, speakers may recognise that lexical entries sharing a common root are related, and an adequate grammar must reflect this fact. This is accounted for in lexicase as a case of lexical derivation. That is, by the triune sign criterion, examples (2) and (4) below contain two different but derivationally related entries *eat/ate*:

<sup>4</sup>Some of Grace's statements support part of this conclusion:

The conceptual elements of the language are the *senses* of its conventional signs. (Grace 1987:89)

A single signans may be paired with more than one sense, and therefore participate in more than one linguistic sign, since each different pairing of a signans with a sense defines a different linguistic sign.

Therefore, what is represented in conventional dictionaries as a single lexical item may have a number of different senses (be a number if (sic) different linguistic signs)... (Grace 1987:88-89)

However, he does not seem to have made the leap to the conclusion that differing distributions also require separate lexical entries.

<sup>5</sup>A grammatically significant environment is one which allows the occurrence of one set of stems but excludes the occurrence of another set of stems. Thus A. and B. are grammatically significant environments, since A allows *load* and *throw* but not *fill*, while B allows *fill* and *load* but not *throw*:

A. X \_\_\_\_\_ Z into Y:      \*filled, loaded, threw  
B. X \_\_\_\_\_ Y with Z:      filled, loaded, \*threw

<sup>6</sup>This view of the nature of the sign has led to the charge that a lexicase grammar will be flooded with homophonous lexical entries. This is however a misconception based on conventions of lexicography: any adequate lexicon must show which word form has which meaning in which environment. In conventional lexicographic practice, these facts are represented (if at all) by lumping together all the distinct entries under the heading of the shared word form (unless the overlap in their respective meanings is subjectively perceived to be too small to justify this), whereas the lexicase view makes the psychologically testable claim that triune signs are separate and independent of each other, a claim for which evidence will be presented below.







(d) Every verb takes a Patient complement (the 'Patient Centrality hypothesis').

(e) AGTs only appear as complements of transitive verbs, and all transitive verbs take an AGT complement.

Case relations in lexicase are primarily syntactic rather than semantic categories. They are set up in order to account for various language-specific and cross-linguistic grammatical regularities, and the minimal semantic content they bear is secondary. For example, PAT is the central participant in an action or state, and while this does play a role in the statement of semantic generalisations, it is much too vague to serve as a defining criterion in recognising the presence of PAT in a given sentence.

More important for the thesis of this paper is the claim that lexicase case roles are perceptual rather than situational. In conventional case grammar and its Government Binding, Lexical Functional Grammar, and other adaptations, case relations/theta-roles refer to situations: an Agent refers to a participant in a language-external situation, and is constant for all sentences referring to that same situation. In lexicase, on the other hand, case relations refer to a speaker's perception of a given situation, and since the same situation can be perceived in different ways, the same external participant may be encoded with different case relations. For example:

- (5) *Mikey ate his mush yesterday.*  
 AGT                    PAT
- (6) *Mikey ate yesterday.*  
 PAT
- (7) *The mush was eaten by Mikey yesterday.*  
 PAT    MNS

This way of defining case roles results in the capture of many nice grammatical generalisations which cannot be formulated in conventional case grammars, and in the avoidance of troublesome paradoxes which bedevil other approaches which are explicit enough to discover them.

### 2.2.1.2 MACRORoles

There is universally one 'macrorole': *actor* (actor). This has more semantic content than the CRs, and can be characterised in notional terms as "the entity to which the action of the verb is attributed", where 'action' is interpreted broadly to include 'actions, happenings, and conditions in general'.<sup>8</sup> This category corresponds in most though not all cases to the definition of Agent in situation-based case grammar or theta-role systems, or 'Initial 1' in Relational Grammar. However, definitions of 'Agent'/'Initial 1' seem to assume that the NP referred to satisfies certain selectional restrictions, and in particular that it be animate, conscious and consenting. Such semantic categories as animacy, intention or consciousness, however, cannot be used as defining criteria in lexicase, which is a syntactic theory. Instead, a syntactic definition of *actor* takes precedence over any such considerations: *actor* corresponds to the AGT of a transitive clause or the PAT of an intransitive clause. Thus some NPs which are analysed as *actor* in lexicase would not match Agent in other theories. For example:

<sup>8</sup>This is Richard Benton's definition of 'Actor'; cf. Starosta (1988:145).

- (8) *The bomb exploded suddenly.*  
actr
- (9) *Bill resembles a wallaby.*  
actr
- (10) *Mustafa knows that he will never be Pope.*  
actr

Moreover, since *actor* too is a perceptual category, it may also match different participants in different perspectives of a single external situation. For example:

- (11) *An asp bit Cleopatra.*  
actr
- (12) *Cleopatra was bitten by an asp.*  
actr

This category, which seems to have no exact counterpart in other theories, plays a crucial role in the statement of certain universal generalisations, as outlined below.

### 2.2.1.3 CASE FORMS

Case forms are grammatical configurations which signal the presence of case relations. Case forms may be realised syntactically (word order, agreement, lexical subclasses such as prepositions or relator nouns) and/or morphologically (case inflections, agreement).

All languages have one case form in common: Nom (Nominative). The nominative case form is the least marked case form in a syntactic as well as a morphological sense. With one limited exception (impersonal verbs), Nom is the case form of the sole NP complement of a one-argument verb. If one NP complement is grammatically required with all verbs while other NPs are optional, the required category is Nom. If only one NP complement agrees with the verb, that NP is in the Nom case form. If one case inflection category is morphologically less marked than the others, that category is Nom. What is referred to as 'grammatical subject' in other frameworks generally matches a case-relation-bearing Nom NP in lexicase, at least for accusative languages.

The case form Acc (Accusative) is universally present in accusative languages. It is the second least marked case form, after Nom. If two NP complements agree with the verb, (as in Swahili, for example), one will be Nom (the one that appears also with one-argument verbs), and the other will be Acc. 'Direct object' in some other frameworks matches a PAT-bearing Accusative NP in lexicase fairly well.

Other case forms are composed of clusters of universal localistic semantic features, especially  $\pm$  sorc (source),  $\pm$  goal,  $\pm$  trmn (terminus), and  $\pm$  ntrr (interior).

## 3. GENERALISATIONS

### 3.1 CAPTURING GENERALISATIONS

If we assume the triune sign position, what does it buy us? A standard criterion for comparing scientific theories which model an unseen reality with observable consequences is economy. Occam's Razor states that theoretical entities should not be multiplied unnecessarily. This could be

viewed as an arbitrary procedure for imposing an arbitrary ranking on alternative theories, but in practice it is thought of by practicing scientists as at least partially empirically motivated. There is an assumption that nature is frugal, and does not do things in complex ways when a simpler means is available; or seen in a less anthropomorphic way, that complex phenomena are most likely the result of the interaction of a limited number of simple fundamental principles. Consequently, a model of some unseen reality which has fewer parts is more likely to faithfully reflect the immanent structure of this reality than one which has more. Based on this reasoning, an account of the properties of the lexicon of a given language is more likely to be faithful to the underlying reality if it is more economical.

I will attempt to show in what follows that the most economical account of the facts of verbal morphology, distribution and semantics is not one which assumes that the semantic properties of verbs (and by extension other parts of speech as well) is a simple mapping of meaning onto external reality, but rather one that assumes that the reality they map onto is psychological rather than physical. A further claim, which I will not attempt to justify here, is that this psychological reality reflected by language is intimately connected with the structure of the language itself, and that in Grace's terms, language constructs its own reality. Thus I am attempting to find evidence from the observable facts of language for part of Grace's arguments against the mapping view of the way languages connect up with reality.

### 3.2 GENERALISATIONS AS EVIDENCE

I have stated some of the assumptions of the lexicase grammatical framework, and indicated some of the ways in which they are similar to the claims made in Grace (1987), as well as some of the ways in which they are different. In the remaining part of the paper, I will present something with which Grace did not concern himself: evidence for my claims drawn from a number of real languages. In such a short paper, it will not be possible to do more than give a sketch of the results of twenty years of lexicase research on more than twoscore and five languages. For details, the reader is referred to Starosta (1988) and the references given there.

My strategy in this section will be to show that the assumptions I have made above about the way languages encode reality, the way the lexicon is organised and stored, and the kinds of grammatical relations available to languages make it possible to capture some nice language-specific and language-universal generalisations. To the extent that more or better generalisations cannot be made in situation-mapping approaches, these generalisations count as confirmations of the lexicase perspective-based analysis and the reality-construction hypothesis which underlies it. In Chomsky's terms, if many observations can be explained in terms of a universal theory, that theory can be said to have achieved *explanatory adequacy*. The implicit or explicit claim here is that other frameworks which make different assumptions (or make no clear assumptions at all) are not able to capture a comparable set.

The examples I will present in this paper are drawn from one area of grammar, verb subcategorisation and verb-to-verb derivation. This phenomenon is related to Grace's claims in two ways: (i) the generalisations that can be extracted from observations of verb classification and clause structure apply neatly across a number of languages from unrelated language families, and are not of the sort that can be motivated by considerations of external environment. Thus they count against Grace's implicit thesis that languages can in principle differ in unpredictable ways. At the same time, (ii) these generalisations must be stated in terms of case relations which are regarded as categories

based on the perception of external situations rather than on categories of external reality per se. Thus they count against the 'mapping' view as well, and lend partial support to the 'reality-construction' point of view.

Lexicase work on case relations and verb subcategorisation over the last twenty years has led to the conclusion that language in general encodes the perception of reality rather than reality as such, and that case relations ('thematic relations') in particular are categories reflecting perception rather than situation: each distinct case frame defining a verb class is a template reflecting a distinct way of perceiving some external reality, a 'perspective' of some external situation. To that extent, such a system is in accord with Grace's idea that language does not map directly onto reality (the 'mapping' view that Grace sees as underlying modern linguistic thought), and is in fact a model of perception rather than a model of the world as such. It differs from Grace's views however in the finding that to a large extent the patterns found in verbal classification and derivation are universal, thus reflecting universals of perception which Grace's ideas would lead us to expect should not exist at all.

#### 4. VERB SUBCATEGORISATION

##### 4.1 GRACE AND LEXICASE

I will begin my discussion of the relation between verb classification and perception by presenting some of George Grace's ideas on this matter and matching them with lexicase ideas.

###### 4.1.1 SENTENCES REPRESENT 'CONCEPTUAL EVENTS'

The constructed realities represented by sentence-level signs will be referred to as *Conceptual Events*. (Grace 1987:31)

The same statement could apply to lexicase.

###### 4.1.2 CONCEPTUAL EVENTS ARE CONSTRUCTED FROM A VOCABULARY OF 'CONCEPTUAL ELEMENTS'

...when we say something, that 'something' consists in large part of what we called a 'conceptual event'. The conceptual event is constructed out of 'conceptual elements' provided by the language and arranged into a structured unit in accordance with the grammatical machinery of the language. Now the conceptual elements out of which the conceptual event is constructed belong to an inventory of conventional elements made available by the language (or the linguistic repertoire). From the standpoint of language, the conceptual events appear as the conventional meanings (which I will call the 'senses') of the conventional linguistic signs of the language (or repertoire). (Grace 1987:75-76)

In lexicase, Grace's 'conceptual elements' correspond most closely to semantic features. Such features represent differential meaning, such that each lexical entry differs from every other entry in at least one feature. The meaning represented is perceptual, and there is no requirement in lexicase that such features have any direct and immutable connection with external objective reality.

Now, an event as represented in a human language is a *structured* unity. It typically consists of a verb and noun phrases or pronouns representing entities of some sort, each

with its own designated role. For example, in our sample sentence ('a man sold the doctor a car') the event consists of an act (of the kind called 'selling') which was performed by an 'agent' (of the kind called 'man') and undergone by an 'undergoer' or 'patient' (of the kind called 'car') directed toward what we may call a 'goal' or a 'referent' (of the kind called 'doctor'). What has been done by the speaker actually is to characterize the event which he/she has abstracted as a particular *kind* of event, in which a particular kind of act (selling) has been performed by a particular kind of agent (a man) on a particular kind of patient (a car), etc. Therefore, not only the unity of the event, but also its structure, is in the eye of the beholder – in this case, the speaker. (Grace 1987:32)

Here the scare quotes around the terms 'agent', 'patient', 'undergoer', and 'goal' suggest that the author does not wish us to take them too seriously. In particular, I think he would object to the suggestion that they are to be taken as categories mapping onto external situations or as formal elements drawn from a universal set. At the same time, the fact that he uses them at all suggests that he expects his readers to understand what he has in mind, and that expectation would seem to presuppose one or both of the above assumptions.

From a lexicase point of view, such terms are conventional colloquial ways of viewing external reality. Other frameworks, such as Fillmorean case grammar, have attempted to build them directly into grammars, but the paradoxes and loss of generalisations that result should be enough reason to make the perpetrators realise that this is a mistake. On the other hand, the lexicase counterparts of these terms, (AGT, PAT etc.), are categories of perception that turn out to have only a rather indirect correlation with them.

#### 4.1.3 LANGUAGE ENCODES CONCEPTUALISATIONS RATHER THAN SITUATIONS

The point which stands out from this discussion is that it is the speaker who chooses what *kind* of event he/she will represent it as being. The choice is not dictated (except in a quite general way...) by what actually happened. However, it is limited, of course, by what I will call the 'conceptual elements' of the language, that is, by the kinds of acts and objects and individuals which the particular language recognises (i.e. has words for). I like to call such a kind of event a *conceptual event* because the event, as an event, is actually a conception of the speaker. It is the speaker who, first of all, identifies the parts of the event which he/she reports as together constituting a single unity and, secondly, decides how to characterize the structure of the event (in this case an act performed by someone, directed to someone, undergone by something) and its various elements (selling as the act, a man, the doctor, a car, respectively, in the other roles). Any given language implicitly recognizes, by means of the lexical and grammatical resources which it makes available, an inventory of possible conceptual events. In the hypothetical speech act which we have been discussing (that is, someone saying, 'a man sold the doctor a car' to report an actual event), we may think of the speaker as having selected one conceptual event from that language's inventory of possible conceptual events. (Grace 1987:33)

To this I would only add that the the verb lexicon must necessarily be the principal repository of these 'conceptual events'.



#### 4.1.4 THE MAPPING BETWEEN CONCEPTUAL EVENTS AND EXTERNAL SITUATIONS IS MANY-TO-ONE

I haven't found a precise statement to this effect in Grace (1987), but given the fact that there is no direct and solid connection between language and external reality, I think it is a fair inference to make. In lexibase it follows from the fact that when a single root occurs in two different distributions, it represents two different lexical entries. The single root may correspond to a single external situation, defined by a constant set of truth conditions, while the distinct entries correspond to distinct perspectives of that situation. So, the mapping is many-to-one.

### 4.2 VERB CLASSES AND PERCEPTION

Lexibase is a type of lexicalist dependency grammar. From the point of view of such a framework, the properties of a construction are determined by the lexical features of the lexical head of that construction. Different clause types encode different perspectives, and each clause type is projected from a different verb type. Each verb type in turn can be seen as a template imposed on reality; in framing an utterance, a speaker must select from among those verb types whose templates are compatible with the external situation he is concerned with. If sentences encode our perceptions of reality, and if the form of the sentence is determined by the lexical properties of the verb, then the verb is the place to look to determine how a language encodes reality.

This means that a description of the verb categories made available by a language will be at the same time a description of the ways which a language makes available for perceiving reality and conveying information about those perceptions. Thus verb classification will not only be the primary content of a description of clause structure, as required by dependency grammar, but will at the same time be at the heart of any characterisation of the way in which a language constructs and/or maps onto reality. The position I will take in this paper is similar to Grace's (and different from Fillmore's, for example) in that I will maintain that the syntactic and semantic subclassification of verbs in every language reflects the perception of reality rather than objective reality itself. However, it differs from Grace's (and is comparable to Chomsky's) in that I will claim that the principal verb categories are defined by a basic set of categories which are comparable and constant across language.

## 5. LEXIBASE GENERALISATIONS

### 5.1 SYNTACTIC GENERALISATIONS

#### 5.1.1 INVENTORY OF VERB CLASSES

Applied to the facts of verbal subcategorisation, the triune sign hypothesis is supported by considerations of economy. If we assume the traditional and Chomskyan binary sign, then the syntactic subclasses of the verbal inventory will be defined in terms of the aggregate of environments in which a verb can occur, and two verbs which do not occur in exactly the same range of environments must be regarded as members of different syntactic classes. On the other hand, the triune sign analysis makes it possible to get by with vastly fewer distinct verb classes, with the same sixteen basic classes turning up in languages as diverse as Swahili and Chinese. The following examples illustrate why this is so:



- (13) *Santa loaded the sleigh with memory chips.*  
 (14) *Santa filled the sleigh with memory chips.*  
 (15) \**Santa threw the sleigh with memory chips.*<sup>9</sup>  
 (16) *Santa loaded memory chips into the sleigh.*  
 (17) \**Santa filled memory chips into the sleigh.*  
 (18) *Santa threw memory chips into the sleigh.*

Conventional linguistic analyses, following traditional lexicographic practice, would say that these data demonstrate the existence of three different syntactically defined verb classes, based on the total range of environments in which the verbs appear. Thus *fill* appears in the environment [\_\_\_NP<sub>1</sub> with NP<sub>2</sub>], *throw* appears in the environment [\_\_\_NP<sub>2</sub> into NP<sub>1</sub>], and *load* appears in both environments:

- Class I. *fill* [+V, +\_\_\_NP<sub>1</sub> with NP<sub>2</sub>]  
 Class II. *throw* [+V, +\_\_\_NP<sub>2</sub> into NP<sub>1</sub>]  
 Class III. *load* [+V, +\_\_\_NP<sub>2</sub> into NP<sub>1</sub>, +\_\_\_NP<sub>1</sub> with NP<sub>2</sub>]

In a triune sign analysis, however, each distinct environment establishes a distinct lexical item, so the *load* in (13) is *load*<sub>1</sub>, a different word from the *load* in (16), *load*<sub>2</sub>. Moreover, *load*<sub>1</sub> belongs to the same syntactic class as *fill* and *load*<sub>2</sub> belongs to the same syntactic class as *throw*, so only two classes are required instead of three:

- Class I. *fill, load*<sub>1</sub> [+V, +\_\_\_NP<sub>1</sub> with NP<sub>2</sub>]  
 Class II. *throw, load*<sub>2</sub> [+V, +\_\_\_NP<sub>2</sub> into NP<sub>1</sub>]

It is easy to see that the maximum number of basic verb classes in the conventional approach will be equal to the total number of *possible combinations* of distinct individual environments, a very large number, while the maximum number in a lexicase grammar will be far smaller, equal to the number of distinct individual environments only. So, the triune sign assumption reduces the inventory of theoretical constructs very significantly.

## 5.1.2 CASE MARKING

### 5.1.2.1 UNIVERSAL

Each distinct environment in which a verb appears defines a different syntactic class, and each such class has its own case-marking pattern, which need only be stated once per class. Thus it is sufficient to indicate which class a given word belongs to, and the case-marking pattern is predictable by general rule. Some of the generalisations extracted in this way turn out to recur across classes and across languages. To cite the most striking:

- (a) Subjects are always marked by the nominative case form in all languages.  
 (b) There are two basic case-marking typologies, and clauses in all languages follow one or the other:  
 (i) **Accusative:** The grammatical subject, as defined by lexicase criteria, signals the presence of the actor macrorole.

<sup>9</sup>Ungrammatical on the reading in which the memory chips are transported to the interior of the sleigh.

- (ii) Ergative: The grammatical subject, as defined by lexicase criteria, signals the presence of the Patient case relation.
- (c) In all languages, ergative or accusative, the case relation of the subject of an intransitive clause is always PAT (Patient).
- (d) The case relation of the subject in a transitive clause in all accusative languages is always AGT (Agent).
- (e) In a transitive clause in all accusative languages, PAT is always marked by the Accusative case form.

Note that these generalisations are not arbitrary decrees imposed on a helpless language; each is justified in terms of other generalisations which it makes possible. Moreover, though it may be objected that the whole system is circular, it is no more circular than theories in physics: in each case, positing invisible hypothetical constructs and interpreting the observations in terms of these constructs makes it possible to explain the regularities and to predict and discover new ones.

- (f) Nom is the highest term in the Comrie case hierarchy (Comrie 1976:263).

The Comrie case hierarchy states an ordering of 'cases' in terms of their priority for grammatical processes such as causativisation and relativisation. Thus subject will always have a higher priority for relativisation than direct object, direct object than indirect object, etc.: SUBJ > DO > IO > ... When 'subject' is defined as Nom according to lexicase criteria, and 'direct object' is defined as Acc-PAT, again in lexicase terms, the hierarchy works perfectly for both ergative and accusative languages. However, if 'subject' in language X is defined in terms of whatever matches the subject of the English translation, as is the usual practice, then the hierarchy will not work for ergative languages.

#### 5.1.2.2 LANGUAGE-SPECIFIC

Consider again the *load/fill/throw* examples given above, but this time with case forms and case relations assigned.

##### 5.1.2.2.1 SITUATIONAL ROLES (FILLMORE/CHOMSKY/BRESNAN STYLE)

- (19) *Santa threw memory chips into the sleigh.*

Objective	Oblique
Patient	Locus

- (20) *Santa filled the sleigh with memory chips.*

Objective	Oblique
Patient	Material

- (21) *Santa loaded memory chips into the sleigh.*

Objective	Oblique
Patient	Locus

- (22) *Santa loaded the sleigh with memory chips.*

Objective	Oblique
Locus	Patient

In situation-based frameworks such as Fillmorean case grammar or Chomskyan theta-theory, the same participant in the event must be given the same case relation/theta-role in all paraphrases. Thus for example *memory chips* must have the same theta-role in (21) and (22). If we choose (21) as basic, then our case-marking rules must somehow allow for the fact that Patient is sometimes marked by Objective case and sometimes by another case (suppose we call it Oblique).

#### 5.1.2.2.2 LEXICASE CASE RELATIONS AND CASE FORMS

(23) *Santa threw memory chips into the sleigh.*

Acc	Lcv	(Lcv = Locative)
PAT	LOC	

(24) *Santa filled the sleigh with memory chips.*

Acc	Abl	(Abl = Ablative)
PAT	MNS	

(25) *Santa loaded<sub>2</sub> memory chips into the sleigh.*

Acc	Lcv
PAT	LOC

(26) *Santa loaded<sub>1</sub> the sleigh with memory chips.*

Acc	Abl
PAT	MNS

Lexicase on the other hand is a perception-based theory, and case relations are not constant across all paraphrases. Thus in examples (24) and (26), the sleigh is perceived as the Patient being affected by having its internal space consumed, while in (23) and (25), the memory chips are perceived as the Patient being affected by being moved through space to a Locus. The case-marking for Patient now only requires one statement instead of two, and in fact it is a universal one: PAT in transitive clauses in accusative languages is marked by the Accusative case, period.

#### 5.1.3 WORD ORDER

Statements about clause-level word order of 'exotic' languages are typically stated in terms of the categories V (Verb), S (Subject), and O (Object), where S is implicitly equated with the subjectively perceived instigator of the action and/or with the subject of the corresponding English gloss, and O (object) is equated with the subjectively perceived undergoer of the action and/or the direct object of the corresponding English gloss.<sup>10</sup> As a result, the word order of the clauses may appear quite variable. For example, such an approach would treat Formosan languages such as Tsou or Atayal (ignoring clitic pronouns) as having variable VOS or VSO word order. Applying the universal lexicase definition of 'subject' (Nom) to the actual language-specific grammatical facts, however, sometimes selects a different NP as the grammatical subject, and results in a new insight: both languages are consistently subject-final, that is, VXS.<sup>11</sup>

<sup>10</sup>Greenberg's universals of word order are a classic (or notorious) example of this procedure.

<sup>11</sup>Where X stands for any constituent other than the main verb or the subject.

## 5.1.4 IMPERATIVES

In imperative constructions across languages, the participant who is being ordered to perform the action is typically pronominalised or left unmentioned. For example:

(27) *(You) get those reindeer off my roof!*

Nom  
AGT  
actr

(28) *(You) get out of here!*

Nom  
PAT  
actr

A typical grammatical description might say that it is the subject which is pronominalised or omitted in such constructions. Note that it is the grammatical subject rather than the 'logical subject' which is relevant here, as shown by passive imperatives:

(29) *(You) be examined by a specialist first!*

Nom	{	'Agent'
PAT		'Logical subject'
actr		'Initial 1'

A lexicase analysis on the other hand would claim that it is the *actor* rather than the grammatical subject which is pronominalised or omitted. Either statement works for accusative languages, but in ergative languages such as the languages of the Philippines and aboriginal Taiwan, *actor* and Nom (grammatical subject) mark different NPs in transitive clauses, and in fact it is the *actor* which is pronominalised, not the grammatical subject:

(30) *Hugasin mo ang pinggan* (Tagalog)

wash	you	the	dishes
+trns	Gen	Nom	
	AGT	PAT	
	actr		

Wash the dishes!

Thus a single lexicase statement works for both language types, while conventional grammars, with only 'grammatical subject' and 'Agent' to work with, cannot make a single statement to cover both situations; if they say the missing NP is the grammatical subject, it won't work for ergative languages, and if they say it is the 'logical subject'/'Agent'/'Initial 1', it won't work for accusative languages because of examples like (29).

## 5.1.5 COREFERENCE

## 5.1.5.1 REFLEXIVES

The situation with reflexives is similar to that with imperatives. Is the antecedent of a reflexive pronoun the grammatical subject or the lexicase *actor*? Both analyses work equally well for accusative languages (with the same exceptions), but the lexicase solution works just as well for ergative languages, while the conventional one does not:

- (31) *Inisip nila ang kanila-ng sarili* (Tagalog; Schachter 1976:503)  
 consider them they self  
 +trns Gen Nom  
 AGT PAT  
 actr  
 They consider themselves.

## 5.1.5.2 INFINITIVAL COMPLEMENTS

Given the lexicase definitions of PAT and *actor*, it is possible to state a universal rule of 'control'. If infinitives are verb forms which exclude the appearance of a subject in the same clause,<sup>12</sup> then the missing subject of an infinitive complement (though not of an infinitival adjunct), that is, PRO in Chomskyan terms, is without exception an *actor* Nom which is coreferential with the PAT in the matrix clause. For example:

- (32) *Mary made John* [ $\Delta$  *get a haircut.*]  
 +trns PAT actr -fint (fint = finite)  
 Nom +trns  
 AGT
- (33) *Mary made John* [ $\Delta$  *fall asleep.*]  
 +trns PAT actr -fint  
 Nom -trns  
 PAT
- (34) *John tried to* [ $\Delta$  *get a haircut.*]  
 PAT -trns actr -fint  
 Nom +trns  
 AGT
- (35) *John tried to* [ $\Delta$  *fall asleep.*]  
 PAT -trns actr -fint  
 Nom -trns  
 PAT

This procedure might appear to pose problems for transitive infinitival complements in ergative languages, since by the lexicase definitions, *actor* will not coincide with Nom in transitive ergative clauses, so the rule will not be able to apply:

- |                           |                                      |
|---------------------------|--------------------------------------|
| (36) Intransitive clauses | Transitive clauses                   |
| a) V NP                   | b) V NP <sub>1</sub> NP <sub>2</sub> |
| -trns Nom                 | +trns Erg Nom                        |
| PAT                       | AGT PAT                              |
| actr                      | actr                                 |

<sup>12</sup>In Chomskyan grammars, infinitives are defined in terms of the absence of tense, but this definition will not work in languages such as Thai, which has no grammatical tense.

- a') V .. NP .. [ V Δ ]  
 PAT -fint actr  
 -trns Nom  
 PAT
- b') \*V .. NP .. [ V Δ NP<sub>2</sub> ]  
 PAT -fint actr Nom  
 +trns AGT PAT
- b'') \*V .. NP .. [ V NP<sub>1</sub> Δ ]  
 PAT -fint actr Nom  
 +trns AGT PAT

So what are the facts? Preliminary results from Ilokano and Atayal indicate that transitive clauses simply do not occur as infinitival complements. Instead, their corresponding antipassives must occur instead, thus meeting the requirements of the control rule:

- c) V<sub>aps</sub> NP<sub>2</sub> NP<sub>1</sub>  
 -trns Obl Nom  
 +aps MNS PAT (aps = antipassive)  
 actr
- c') V .. NP .. [ V NP<sub>2</sub> Δ ]  
 PAT -fint Obl actr  
 -trns MNS Nom  
 +aps PAT

The following examples from Ilokano (Clausen 1986, lexicase analyses) illustrate this phenomenon:

- (37) \*Mabuteng dagiti ubbing nga luganan Δ diay roller coaster  
 be-afraid the children to ride-on Gen Nom  
 PAT +trns AGT PAT  
 actr

The children are afraid to ride on the roller coaster.

- (38) Masadsadot -ak nga ag-luto iti lumpia Δ  
 be-too-lazy I to cook spring  
 PAT -fint rolls Nom  
 -trns PAT  
 +aps actr

I am too lazy to cook lumpia.

If it turns out to be generally true that transitive ergative clauses cannot be infinitival complements, conventional accounts could only stipulate that as an apparently arbitrary fact. In lexicase, however, it is a straightforward consequence of the control rule: the notional 'agent' of a transitive clause (NP<sub>1</sub> in example (36b) above) must be reinterpreted as a PAT in order to bear the features *actor* and *Nom*, and by universal lexicase definition, this is exactly what antipassive derivation does.





## 5.2.2 SEMANTIC DIFFERENCES AMONG DIFFERENT USES OF THE SAME ROOT

There are often semantic differences among different uses of the same root, depending on which of the participants in the external situation is viewed as the central participant, that is, the Patient. For example:

(44) *John is drinking<sub>1</sub> kaoliang wine lately.*

AGT +trns PAT  
actr

(45) *John is drinking<sub>2</sub> lately.*

PAT -trns  
actr

Example (44) describes a transitive action, in which an actor-Agent affects a Patient. In (45), on the other hand, an intransitive action is depicted. The perspective is of the actor-Patient, *John*, carrying out an activity, with the undergoer of the activity treated as irrelevant.

This distinction is shown lexicasually by treating the two *drinking's* as different lexical items belonging to different verb classes. *Drink<sub>1</sub>* belongs to the transitive two-argument total affect class with verbs like *devour* and *create*, while *drink<sub>2</sub>* belongs to the intransitive activity class with verbs like *giggle* and *dine*. Thus in (44), *kaoliang wine* is the PAT, the affected entity with an affect verb, while in (45), *John* is the PAT, the actor of an activity verb. The external objective situations might be the same, but the two sentences view them differently, and this is reflected in a different assignment of case relations.

With verb classes defined in terms of perspective, it is possible to make uniform and exceptionless generalisations about entire syntactically defined classes; in the above examples, PAT is the affected entity in all transitive affect verbs. Similarly, if the same root appears in the *devour* and in the *dine* class, a set of parallel semantic differences will be found between the two uses and between *drink<sub>1</sub>* and *drink<sub>2</sub>*, *devour* and *dine*, etc. Formally, this will be accounted for in terms of a lexical derivation rule which introduces the syntactic and semantic differences in the process of derivation.

## 5.2.2.1 CENTRALITY

The Patient Centrality hypothesis (all verbs take a Patient in their case frames) is both a syntactic and a psychological claim: PAT is the perceptual centre of every predication. This is frequently reflected in aspectual differences, as in the following pairs of classic examples taken from the Fillmorean case grammar literature and provided with partial lexicase analyses:

(46) *John sprayed paint on the wall.*

+trns PAT LOC

+lctn

(lctn = location)

+motn

(motn = motion)

(47) *John sprayed the wall with paint.*

+trns PAT MNS

+afct

(afct = affect)

(48) *Bees are swarming in the garden.*

PAT      -trns                      LOC  
             +lctn  
             -motn

(49) *The garden is swarming with bees.*

PAT              -trns                      MNS  
                     +afct

In examples (46) and (48), the verbs are location verbs (+lctn), and the LOC states the location of the PAT, *paint* or *bees* respectively, while in (47) and (49), the verbs are affect verbs and state an affect on the PAT, *wall* and *garden* respectively. It has often been noted that in (47) and (49)-type sentences, the action or state tends to be viewed as total; thus the wall is covered with paint in (47), and the garden is full of bees in (49). This is explained in lexicase as a result of the class interpretation of PATs of affect verbs as being totally affected. In an analysis which treats the *sprays* and *swarms* as the same lexical item, however, this explanation is not available. And in fact the lexicase explanation is able to treat the transitive and intransitive examples as involving the identical 'affected Patient' analysis only because of the Patient Centrality hypothesis, which requires the subjects of intransitive verbs such as (48) and (49) to be analysed as PAT, rather than as Agent or Locus as would be the case in Fillmore-type case grammars.

#### 5.2.2.2 ASPECT

This same analysis then helps to explain related phenomena, for example, the scope of adverbs:

(50) *John loaded<sub>1</sub> the wagon with hay.*

+afct              PAT

(51) *John loaded<sub>2</sub> the hay on a wagon.*

+lctn              PAT  
             +motn

Although these may seem to mean the same thing, the difference becomes obvious when we add an adverb such as *completely*, or a matrix verb such as *finish*:

(52) *John completely loaded<sub>1</sub> the wagon with hay.*

+afct              PAT

(53) *John completely loaded<sub>2</sub> the hay on a wagon.*

+lctn              PAT  
             +motn

(54) *John finished loading<sub>1</sub> the wagon with hay.*

+afct              PAT

(55) *John finished loading<sub>2</sub> the hay on a wagon.*

+lctn              PAT  
             +motn

With examples (52) and (54), the PAT of the affect verb is completely affected, that is, the wagon is full of hay, whereas with (53) and (55), the PAT of the transportation verb is completely transported to the LOC, that is, all the hay has been transported to a wagon.

## 5.3 MORPHOLOGICAL GENERALISATIONS

## 5.3.1 ZERO AND NON-ZERO DERIVATION

The relation between load<sub>1</sub> and load<sub>2</sub> in the above examples and in many similar examples is treated in lexicase as a matter of lexical derivation. Since no derivational morphology is involved, some linguists have been reluctant to accept such analyses. However, the widespread use of zero derivation is simply an idiosyncratic fact about English,<sup>14</sup> and the expected morphology frequently turns up with analogous derivation processes in other languages, sometimes closely related ones such as German:

- (56) *The bus travels<sub>1</sub> on this street.*  
 PAT -trns LOC
- (57) *The bus travels<sub>2</sub> this street every morning.*  
 AGT +trns PAT
- (58) *Der Bus faehrt auf dieser Strasse. (=56))*  
 PAT -trns LOC
- (59) *Der Bus befahrt diese Strasse nicht.*  
 AGT +trns PAT  
 The bus doesn't take this street.

An analysis in terms of lexical derivation then not only accounts for the syntactic and semantic differences in such examples (the fact that in the second example of each pair, the location has the grammatical appearance of a direct object, and the accompanying perception that the location in these examples is in some sense affected), but also accounts for the morphological differences in German that go with such differences in distribution and meaning.

A similar situation can be seen in the derivation of intransitive verbs from transitive verbs. For example, many linguists would attribute the difference between (60) and (61) to the fact that the verb 'to eat' takes an optional direct object:

- (60) *Hugo eats mustard cabbage.*  
 (61) *Hugo eats.*

Such an analysis is not able to account for the fact that other semantically similar verbs may either require the presence of a direct object (*devour*) or reject direct objects (*dine*). Such lexical gaps support the lexicase claim that two different and independent verb stems are involved here, *eat*<sub>1</sub> and *eat*<sub>2</sub>. These stems are related by a rule of intransitive derivation which drops the original Patient and reinterprets the old Agent as Patient to satisfy the requirements of the Patient Centrality hypothesis:

- (62) *Hugo eats<sub>1</sub> mustard cabbage.*  
 Nom +trns Acc  
 AGT PAT  
 actr

<sup>14</sup>And one which is destined to bring disaster to morphological theories such as Lieber's (1981), in which the properties of derived words are located in the lexical entries of the derivational affixes themselves. Lexicase was never lured into this dead-end street because in lexicase, affixes have no independent existence in the lexicon.

- (63) *Hugo eats*<sub>2</sub>.  
 Nom -trns  
 PAT  
 actr

This analysis receives further indirect support from other languages in which functionally identical object omission is morphologically marked. Thus in Sora, the corresponding verbs would be morphologically different, since detransitivised verbs take an *-n* suffix. For example:

- (64) *Tabono kAmbUnjElAn jUmte*  
 Nom Acc +trns  
 AGT PAT  
 actr  
 Tabono eats pork.

- (65) *Tabono jUmten*  
 Nom -trns  
 PAT  
 actr  
 Tabono eats.

In ergative languages, the difference would be even more dramatic. Thus in Philippine and Formosan languages, the PAT of the transitive sentences (cf. *mustard cabbage* in example (62)) would be in the Nominative case form, but when the original Patient is downgraded or omitted in the intransitivisation process<sup>15</sup> and the original Agent (cf. *Hugo*) is reinterpreted as Patient, the new Patient takes on the Nominative case form, so there is a change in case form as well as number of arguments. In addition, the application of this process is marked by reflexes of the Proto Austronesian prefix \*MU. Here there can hardly be any question that a new word form is involved, and it is only the fact that English is an accusative language with a proclivity for zero derivation that raises any doubts about the corresponding situation in English. Once again, then, the triune sign analysis makes it possible to recognise the same process at work in different languages and capture it with a single basic rule of intransitivisation.

### 5.3.2 AUSTRONESIAN 'FOCUS'

The correlation between syntactic, semantic, and morphological facts in a lexicase analysis of verbal derivation is most dramatically reflected in the lexicase analysis of Western Austronesian and Micronesian focus systems. From a lexicase analysis, what has been called 'focus' is simply a verbal derivation system in which different case relations are reinterpreted as PAT, with the expected syntactic, semantic and morphological differences that might be expected in such a process (cf. Starosta 1986).

It is rather gratifying that the object focus systems of Micronesian languages and the subject focus systems of Philippine and Formosan languages then turn out to be functionally as well as etymologically identical, with nothing but a rather superficial ergative-to-accusative shift accounting for the seemingly dramatic differences (cf. Starosta, Pawley and Reid 1982).

<sup>15</sup>The forms derived by this process and/or the constructions in which they occur are referred to as 'antipassive' in ergative languages.

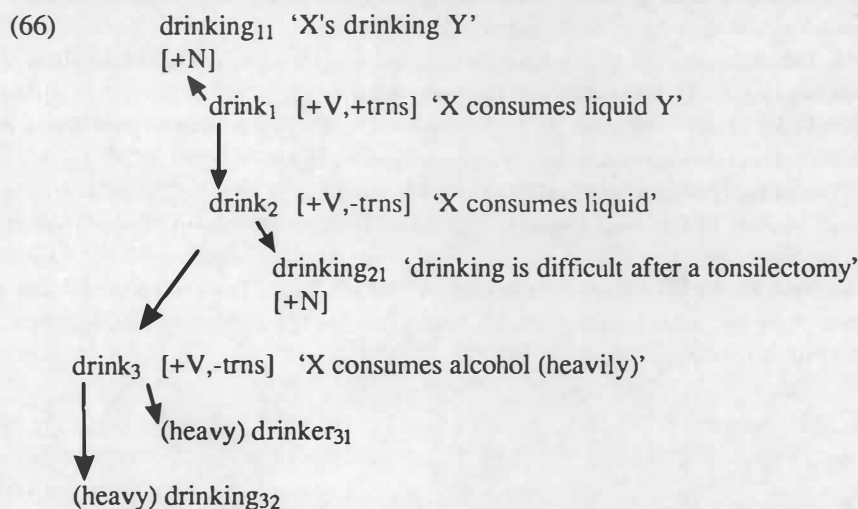


## 5.4 PROPERTIES OF DERIVATION

To sum up, the lexicase account of verb classification and verb root distribution as lexical derivation is plausible for a number of reasons. By this account, the same root can be derived into different classes. When it is, it takes on the semantic properties of that class (it changes its 'perspective'), even though it may still refer to more or less the same external situation. The consistent differences introduced in this change of perspective can be formalised in a straightforward way in terms of standard lexicase derivation rules, resulting in a systematic account of the relationship among the different uses of a verb root in terms of the perspective imposed by membership in each distinct verb class. This analysis treats the English verb classification as just a special but regular case of the kinds of morphologically marked verbal derivation systems found in German, Munda, Salish or Austronesian.

Because derivation is by its nature not productive, the gaps in the derivational 'paradigms' and the variation among speakers is no longer an embarrassment (cf. Salkoff 1983) but strong support for a derivational rather than a transformational analysis. Derivation forms 'word families' rather than regular paradigms. It applies sporadically, when there is a need for it, and since each speaker has his own lexicon and each word has its own history, the patterns will vary from speaker to speaker.

As with other cases of derivation (e.g. *\*ruly*: *unruly*), the derivative can continue to exist after the source has been lost, and it can change independently of other derivationally related items. This can be seen in nominalisations such as (*heavy*) *drinker*, resulting from the following scenario:



*Drink*<sub>2</sub> and *drink*<sub>3</sub> are distinct entries; *drink*<sub>3</sub> implies the consumption of alcoholic beverages and idiosyncratically takes the adverb *heavily*, while *drink*<sub>2</sub> does not. *Drink*<sub>2</sub> serves as the derivational source of the noun *drinker*<sub>31</sub>, which inherits the license to take the adjective *heavy*. Similar examples can be constructed for *eat*, *bite* or *smoke*, but not every set of forms will have exactly the same number of elements in the same relationships to each other, because we are dealing with lexical derivation, and lexical derivation is not productive. In addition, we should not be surprised to find that individual speakers differ on the existence or meanings of individual forms: these forms are separate lexical items, and the lexical inventories of individual speakers differ.

## 6. COMPARISONS

In conclusion, we have found that the lexicase position on the connection between language and reality is similar to George Grace's in stating that language does not map directly onto reality, but rather reflects psychological constructs. It is different from Grace's conception in that it takes no position at this point as to how far the reality which speakers perceive is determined by the structures of their respective languages. The lexicase position comes closer to Chomskyan doctrine in finding that many of the perceptually founded grammatical elements which lexicase posits can be found across a very broad range of human languages, and unlike Grace's explication, it supports this claim by citing as evidence a list of linguistic generalisations that are made possible by assuming the existence of these elements.

## BIBLIOGRAPHY

- CLAUSEN, **Josepha Paz**, 1986, Sentential complements in Ilokano. MA thesis, Department of Linguistics, University of Hawaii.
- COMRIE, **Bernard**, 1976, The syntax of causative constructions: cross-language similarities and divergences. In M. Shibatani, ed. *Syntax and semantics, 6: The grammar of causative constructions*, 261-312. New York: Academic Press.
- GRACE, **George W.**, 1987, *The linguistic construction of reality*. London: Croom Helm.
- LIEBER, **Rochelle**, 1981, *On the organization of the lexicon*. Bloomington: Indiana University Linguistics Club.
- SALKOFF, **M.**, 1983, Bees are swarming in the garden. *Language* 59:288-346.
- SCHACHTER, **Paul**, 1976, The subject in Philippine languages: topic, actor, actor-topic, or none of the above? In Charles N. Li, ed. *Subject and topic*, 491-518. New York: Academic Press.
- STAROSTA, **Stanley**, 1986, Focus as recentralization. In Paul Geraghty, Lois Carrington, and Stephen Wurm, eds *FOCAL I: papers from the Fourth International Conference on Austronesian Linguistics*, 73-95. PL, C-93.
- 1988, *The case for lexicase: an outline of lexicase grammatical theory*. London: Pinter.
- forthcoming, A grammatical typology of Formosan languages. Fang-kuci Li memorial volume. Taipei: Institute of History and Philology, Academia Sinica.
- STAROSTA, **Stanley**, **Andrew K. PAWLEY** and **Lawrence A. REID**, 1982, The evolution of focus in Austronesian. In Stephen A. Wurm and Lois Carrington, eds *Papers from the Third International Conference on Austronesian Linguistics*, vol.2: *Tracking the travellers*, 145-170. PL, C-75.

## TWO SYMMETRICAL VERBAL ASPECTS IN OCEANIA

CLAUDE TCHEKHOFF

### 1. INTRODUCTION

In this paper I shall compare the perfective verbal aspect in Polynesian Tongan (Tchekhoff 1973, 1978b, 1979, 1987), and the so-called 'antipassive' (Tchekhoff 1985) in Yandruwandha (Breen 1975) and Diyari (Austin 1981). The latter are two Australian languages from the southern part of the Northern Territory. The two groups show formal similarity and a striking mirror-image for meaningful contents. In turn, I shall define my view of aspect in general, then go into Tongan examples and finally into Yandruwandha and Diyari.

Aspect has been conclusively analysed elsewhere (Comrie 1976), but it can be brought down to two extralinguistic universals: for every living creature, two parameters are of the essence, before anything else: they are time and space. These universals are necessary and absolutely central to life. Small wonder that they find their way into man's thinking and his expression of this thought. And there, in man's spoken expression, they are just as central. The onus of expressing these two central universals, consequently falls on the central unit of a spoken utterance, and that is the verb.

Verbal aspect is the translation of these basic facts into language. It expresses the behaviour of the verbal operation in regard to time – for instance an event or state can take place over a prolonged period – that is the progressive or inchoative aspect; or it can be expressed punctually, outside of time – such is the aorist of classical Greek; or again it can be shown to be complete, from a successful start to an effective impact on its goal – this can be called the perfective aspect. The behaviour of the verbal operation then implies both its beginning and its end, its two terms, start, *terminus a quo*, and goal, *terminus ad quem*. Tongan examples will illustrate this definition of the perfective aspect. Its opposite would express an unsuccessful verbal operation, one that fails, that does not reach its goal in space. The space parameter is thus relevant to aspect, just as time is.

A consequence of this fact is that aspect thus extends beyond the verb; it also concerns both (pro)nominal participants to the verbal operation. Aspect can thus be said to cover the whole sentence.

Basic utterances in Tongan and also in Yandruwandha and Diyari have an ergative morphology with nominal participants (Dixon 1979, Tchekhoff 1973, 1978a, 1979 etc.). Tongan has an accusative construction with pronouns. Let us compare the two constructions:

ERGATIVE: the one and only participant of an intransitive verb (Vi) (henceforth Si), like the patient (P) of a transitive verb (Vtr) (henceforth Ot), is in the ABSOLUTIVE case. Whereas the agent (A) is marked as such (henceforth St) because in translation it corresponds to the transitive subject of an ACCUSATIVE construction.

In an ACCUSATIVE construction, it is the other way around: Si and St are in the same zero unmarked, NOMINATIVE case, and Ot is marked.

Let us now illustrate these statements through some examples in Tongan first:

- (1) ACCUSATIVE     'Oku   ou   'alu.  
 PRES<sup>1</sup> Si:I   Vi:go  
 I am going.  
 'Oku   ou   ui       koe.  
 PRES St:I   Vtr:call   Ot:you  
 I call you.
- (2) ERGATIVE       'Oku   'alu   'a Tolu.  
 PRES go   Si  
 Tolu is going.  
 'Oku   ui   'e Fifita 'a Tolu.  
 PRES call   ERG.A   Ot  
 Fifita calls Tolu.

## 2. PERFECTIVE ASPECT IN TONGAN

In order to have the status of full morphemes, with both meaning and form, aspect and tense should be able to commute independently: for example 'I worked', past tense, general aspect; 'I was working', same tense, progressive aspect. But: 'I am working', same progressive aspect, different tense. It is thus clear that here tense and aspect are two different morphemes. They commute independently on the same hierarchical basis. This is the case for the perfective aspect in Tongan. Here aspect and tense follow different syntactic paradigms.

Tongan tenses are four: present *'oku*, distant past *na'e*, future *'e/te*, perfect or close past extending into the present *kuo* (Tchekhoff 1979 section 3.20).

An *-i* suffix marks the perfective aspect. Compare (3) unmarked for aspect, and (4) marked for perfective aspect hereafter:

- (3) *Na'e fana 'e Tolu 'a Fifita.*  
 PAST shoot St.A    Ot.P  
 Tolu shot at Fifita.

But in this sentence nothing is said of the outcome of the action, whether it was successful or not.

<sup>1</sup>Abbreviations used are as follows: ABS – absolutive; ALL – allative; APS – antipassive; AUX – auxiliary; CLP – perfect, close past extending into the present; CONT – continuous; DAT – dative; DEM – demonstrative; EMPH – emphatic; ERG – ergative; FUT – future; INS – instrument; LOC – locative; NOM – nominative; PERF – perfective aspect; PPT – past/present participle; PRES – present; SUF – suffix.

- (4) *Na'e fana'i 'e Tolu 'a Fifita.*  
Tolu shot down Fifita.

Each one of my three informants had a slightly different use of the *-i* verbal suffix. But each one of them illustrated a facet of its same general meaning: Tevita, Siosiane and Pauli (her husband) showed complementary uses of this aspect, and for all three, it answered to the following definition: "With a perfective aspect, the verbal operation is expressed in its entirety, its way of encompassing Time, and also including both its terms, start and finish, in Space" (Tchekhoff 1979:238).

For Tevita, those two terms (participants) are included in a verb marked for the perfective aspect, whether explicit or not: if one of them is not expressed, it is nonetheless understood, and will be translated as 'someone specific but whom we do not want to name':

- (5) *Na'e fana'i 'a Fifita.*  
Someone specific whom we are not naming has shot down Fifita.
- (6) *Na'e fana'i 'e Tolu.*  
Tolu has shot down someone we are not naming.

Example (6) is a perfectly grammatical sentence in Tongan, whereas one unmarked for aspect and with the agent but not the patient expressed is unacceptable:

- (7) *\*Na'e fana 'e Tolu...*  
PAST shoot A  
Tolu shot at...

This shows the virtual presence of the patient in example (6) above, and explains why (6) is acceptable, but (7) is not.

One last remark on Tevita's use of the perfective aspect in Tongan: it can go with any tense, present, future and both past tenses. This shows the complete independence of aspect in regard to tense. A verbal operation may not yet have taken place, and be perfective: this simply means that it is envisaged in its entirety, whatever its tense. In Tevita's speech, there are no constraints as to the use of any tense with the perfective aspect.

Such is not the case with Siosiane: in her use of the perfective aspect, it is compatible only with the two past tenses:

- (8) *Na'e/kuo hiva'i 'e he kau hiva 'ae himí.*  
PAST/CLP sing.PERF A the choir the hymn  
The choir sang the hymn.

But:

- (9) *\*'Oku'e hiva'i 'e he kau hiva 'ae himí.*  
PRES/FUT sing.PERF A the choir the hymn

is unacceptable.

This state of things obtains in Siosiane's speech, because for her a perfective verb means that the operation is accomplished, over. Aspect thus shows a closer relationship to tense than it does in Tevita's idiolect.

Besides this point, there are in Siosiane's speech two other constraints which concur to show that a perfective verbal operation 'hits the mark' so to speak: its two (pro)nominal participants must be

specific and singular. If the agent is successful in its operation, if it reaches its mark, the patient evidently must be attainable, it must not be spread out. Conversely an aspectually general unmarked verb will accept neither specific nor singular participants. Here the patient must be general, non-specific.

These constraints again go to show that aspect concerns the whole sentence, agent and patient as well as the verb. Here are some examples:

Plural agent and patient, general unmarked verb:

- (10) *Na'e tokonia 'e he kakai 'ae tamaikí.*  
 PAST help A the people the children  
 The people helped the children.

But a collective agent is incompatible with a marked perfective verb:

- (11) \**Na'e tokonia'i au 'e he kakaf.*  
 PAST help.PERF Ot:me A the people  
 The people helped me.

For this sentence to be correct, the *-i* suffix must be left out. With it, the agent must be singular and specific. He must be completely and precisely implicated in the perfective verbal operation:

- (12) *Na'e tokonia'i au 'e Sione.*  
 PAST help.PERF Ot:me A John  
 John helped me.

The case is the same for the patient of a perfective verb:

- (13) \**Na'e fana'i 'e he tamaiki 'ae manupuná.*  
 shoot down A the children the birds  
 The children shot down the birds.

Siosiane refused to accept example (13) not only for its collective agent, but for its patient too: 'because', said she, 'you cannot shoot down all the birds there are'. In short, Siosiane will use the perfective suffix *-i* with singular and specific nominal participants only.

However, there are certain verbs where the *-i* suffix has a predominant sense of completeness, of strength, rather than a grammatical value as an aspect. In such cases, Siosiane accepts these verbs + *-i*, in all tenses:

- (14) *'Oku'e ne ta'i au.*  
 PRES/FUT St:he beat up Ot:me  
 He beats/will beat me up.

In Pauli's idiolect, the *-i* suffix is incompatible with the present and future tenses, just as it is for Siosiane. But it no longer bears any constraints as to the nominal participants in the sentence. Its meaning is simply that of a completed, intensive, verbal operation, just as that of Indonesian *-i* (Tchekhoff 1978b).



## 3. THE PERFECTIVE SUFFIX WITH INTRANSITIVE VERBS

Here again, the effect of an *-i* suffix on an intransitive verb, is to make it transitive, just as in Indonesian (Tchekhoff 1978b), since it now implies its patient, whether expressed or not (example (6) above):

- (15) *Na'e mohe 'a Sione.*  
 PAST Vi:sleep ABS John  
 John slept.

But:

- (16) *Na'e mohe 'i 'e Sione 'ae 'aho kakatoá.*  
 PAST sleep.PERF A John the day entire  
 John slept the whole day away.

It would be interesting to study the perfective aspect with complex irrealis, interrogative, negative and prohibitive sentences. Unfortunately this study has not yet been possible. It will be treated elsewhere in the future. It would then be possible to pinpoint the limits of the symmetry between the perfective aspect as in Tongan, and its contrary, also positively marked, the imperfective aspect, sometimes named 'antipassive' (Tchekhoff 1985:372), as in some Australian languages (Tchekhoff 1985, 1987).

Let us now consider the so-called 'antipassive' (Silverstein 1976, Tchekhoff 1985).

The great majority of Australian Aboriginal languages have an ERGATIVE morphology, with a marked agent (henceforth St) in the ERGATIVE case, and an unmarked patient (henceforth Ot) in the zero or ABSOLUTE case. Here is an example in Diyari (All Diyari examples cited here are taken from Austin 1981), spoken in the southern part of the Northern Territory:

- (17) *N̄aṯu n̄aṇa wiḷa kaḷka-yi.*  
 ERG.A. ABS.3SG.DEM woman wait-AUX  
 I am waiting for the woman.

The next example shows the same sentence, but the verb is marked with the 'antipassive' suffix. It becomes intransitive.

- (18) *N̄aṇi kaḷka-tadi-yi n̄aṇkaṇu wiḷa-ni.*  
 Si wait-APS-AUX 3SG.DEM woman-LOC  
 I happen to be waiting for the woman.

In example (17), the operation is wilful; in example (18), it is more desultory, or has happened accidentally. Again:

- (19) *N̄aṯu yinaṇa ḍanka-ṇa wara-yi.*  
 St:I Ot:you find.PPT PRES-AUX  
 I have found you (after looking for you).
- (20) *N̄aṇi ḍanka-tadi-ṇa wara-yi yiṅkaṇu.*  
 Si Vi:find.APS.PPT PRES-AUX 2SG.DAT  
 I happened to bump into you by chance.

Note that the differences between examples (17) and (19) on the one hand and (18) and (20) on the other, are as follows: in the second pair, the action of the verb-cum-suffix- is less wilful. Hence Si and not St, the subject is not an agent. Also its impact at the other end is less direct, hence the

intransitive verb and oblique LOCATIVE and DATIVE cases of the erstwhile patients. If one compares this to the perfective examples in Tongan, (examples (15) and (16) above), it is obvious that in one case as in the other, the verb form commands the cases of subject and object. With the perfective aspect and the 'antipassive' both, the influence of the suffix thus extends to the whole sentence.

The next two examples are Yandruwandha. They also illustrate the accidental feature of the 'antipassive' suffix. Yandruwandha is now extinct. It was spoken in the southern part of the Northern Territory, just as Diyari was, but this implies no particular connection between the two languages. (All Yandruwandha examples cited here are taken from Breen 1975):<sup>2</sup>

- (21) *Njapali yulpurruli ŋunu nyunma-indri-ŋa na.*  
 ERG:flood Si.NOM he CLP:drown CLP  
 He drowned in the flood.

The accidental feature of an 'antipassive' verb can be hardened, and made to express unsuccessful operations, as in the following derivation: *ŋanma* meaning 'to hold fast'. It is used with the same suffix *-indri-* with various meanings according to context:

- (22) *ŋanma ŋanma-indri-* to stutter  
 to be hard to follow (when speaking)  
 to limp  
 to balk (speaking of a horse), etc.

The common point for all these uses is that they all mean an unsuccessful, negative operation derived from a positive verb base.

A verb with such a suffix can show a direct object, although most of the time, it becomes intransitive. But if it does take an object, the latter is general, non-specific. Breen has felicitously named these formations 'pseudo-transitive':

- (23) *Muŋuru ŋanyi tayi-indri-na ŋala walya ŋu ŋanyi yaŋu.*  
 ABS:tucker Si:I eat-SUF-CLP not yet Si:I full  
 I have eaten but I am not yet full.

Diyari:

- (24) *Njandi puka wayi-ŋadi-ŋa ŋama-yi.*  
 ABS.mother ABS:food cook-SUF-PPT sit-AUX  
 Mother is sitting down cooking food.

Since the patient of a verb with this suffix is not specific, it follows that the formation can be used to express repetitive operations.

Yandruwandha:

- (25) *Tayi-indri-ŋa ŋina ŋangu-ŋaŋalatyi.*  
 eat-SUF-FUT he stay:CONT.EMPH  
 He could eat when he stopped there continuously.

<sup>2</sup>In his analysis of the Yandruwandha suffix *-indri-* Breen does not use the denomination 'antipassive'. Therefore, I have not done so here either in my examples taken from his manuscript. But Austin does analyse *-ŋadi-* as an 'antipassive' suffix. The uses of the two are exactly parallel from one language to the other.

Exactly symmetrical and contrary to what we have seen for the Tongan perfective aspect (example (12) above), and because of its lack of impact on a putative patient, an 'antipassive' verb can express a reflexive and a middle operation, that is, doing something for oneself (Benveniste 1966). The action of the verb reverts to its subject, rather than extending outwards. Here are two Yandruwandha examples:

(26) *Tawapaṅḍila ḡanyi kaririḡari ḡapa maṅdra-indri-ḡa.*  
 PRES:go down Si:I ALL:river ABS:water scoop up-SUF-FUT  
 I am going down to the river to scoop up some water (for myself).

(27) *ḡapaḷa ḡulu wani ḡraḡa-indri-ḡaḡa.*  
 well he corroborree sing-SUF-CONT  
 Well, he was singing to himself.

Reciprocal being reflexive in the plural, in Yandruwandha the same suffix is used here again:

(28) *Maḡramityi -li ḡanaru ḡranyi-indri-ḷa.*  
 stone eye INS they hit-SUF-PRES  
 They are throwing stones at each other.

To sum up, with a perfective aspectual suffix, an intransitive verb becomes transitive. Its patient must be definite and specific (example (12) above). Its agent operates wilfully. With the 'antipassive' suffix, the verb itself can become intransitive, the patient is demoted to an oblique case, and the agent to a mere Si (example (20) above). If the verb remains transitive as it can do, the verbal operation then applies to a general, non-specific object. As will be remembered, a verb with a perfective aspect (example (4) above) must hit the mark. It has a direct impact on its patient. Contrarily, the lack of impact of the verbal operation on a putative object commands the use of an 'antipassive' verb to express unsuccessful operations, with either no resulting impact or a faulty one. It can thus be seen that the meaning and functional mechanism of the 'antipassive' are the exact mirror-image of the perfective aspect, while both are expressed by a verbal suffix.

For a proper methodological approach, it is essential to note that the comparison is made licit and permitted because each language has within its own system oppositional pairs with and without suffix. The oppositions thus obtain within each system. After their existence has been established independently inside each language, one system can then be compared to the other. It is therefore closer to factual linguistic reality to give the 'antipassive' its real functional value as an imperfective aspect (Tchekhoff 1985:372).

This interpretation can be corroborated by comparison with other, entirely different language-systems: I have done so for Russian and some Caucasian languages (Tchekhoff 1985). In all these languages the same meaningful contents are expressed, under very different guises; but they always have to do with the entire sentence. This approach can iron out some syntactic traits which otherwise would remain unexplained (1985, section 4).

## BIBLIOGRAPHY

- AUSTIN, Peter, 1981, *A grammar of Diyari, South Australia*. London and New York: Cambridge University Press.  
 BENVENISTE, Emile, 1966, Actif et moyen dans le verbe. In *Problèmes de Linguistique générale* 1:168-175. Paris Gallimard.  
 BREEN, Gavan, 1975, Innamincka talk. (The Innamincka dialect of Yandruwandha). MS, preliminary version. A.I.A.T.S.I.S. Library.

- COMRIE, Bernard, 1976, *Aspect*. London and New York: Cambridge University Press.
- DIXON, R.M.W., 1979, Ergativity. *Language* 55:59-138.
- SILVERSTEIN, Michael, 1976, Hierarchy of features and ergativity. In R.M.W. Dixon, ed. *Grammatical categories in Australian languages*: 112-171. Canberra: A.I.A.S.
- TCHEKHOFF, Claude, 1973, Parataxe et construction ergative avec exemples en avar et tongien. *Bulletin de la Société de Linguistique de Paris* 68/1:255-268. Paris: Klincksieck.
- 1978a, *Aux fondements de la syntaxe: l'ergatif*. Paris: Presses Universitaires de France.
- 1978b, Typology and genetics: some syntactic conclusions that can be drawn from a functional comparison between Indonesian verbal suffix *-i* and Tongan *-i*. In S.A. Wurm and Lois Carrington, eds *Second International Conference on Austronesian Linguistics: proceedings*, 367-382. PL, C-61.
- 1979, *La construction ergative en avar et en tongien*. Publications de la Sorbonne. Paris: Klincksieck.
- 1985, Aspect, transitivity, "antipassives" and some Australian languages. In Frans Plank, ed. *Relational typology*, 359-390. The Hague: Mouton.
- 1987, "Antipassif" aspect imperfectif et autonomie du sujet. BSL 82/1:43-67. Paris: Klincksieck.

## WANEM BISLAMA?

DARRELL TRYON

### 1. INTRODUCTION

Much of the current debate about the genesis and development of Pacific pidgins centres around the recent work of Peter Mühlhäusler and Roger Keesing, although to be fair one should not forget the contributions of other scholars such as Bickerton (1975, 1981), Charpentier (1979), Charpentier and Tryon (1989), Clark (1979), Crowley (1989), (1990), Dutton (1985), Jourdan (1985), Mosel (1980), Siegel (1987) and Tryon (1987, in press).

We will first look at the substance of the debate which is currently exercising the minds of Keesing and Mühlhäusler. In 1986 Mühlhäusler published his *Pidgin and Creole linguistics*, while in 1988 Keesing replied with his *Melanesian Pidgin and the Oceanic substrate*. Since that time, Mühlhäusler has reviewed Keesing (Mühlhäusler 1989) and Keesing has replied to a number of critics and reviewers, namely Belikov, Bickerton, Mühlhäusler, Romaine and Siegel (Keesing, in press).

Mühlhäusler maintained, inter alia, that the principal features of Melanesian pidgins are attributable to language universal forces, with little substrate influence, while Keesing takes the view that Melanesian pidgins are largely calques based on the morpho-syntax common to Oceanic Austronesian languages. Although both Mühlhäusler and Keesing make a large number of claims about the development of Pacific pidgins, one significant gap in the linguistic evidence upon which they base their arguments is the almost total absence of the evidence which may be adduced from Vanuatu Bislama, both current and historical. Indeed, Keesing (1988:vi) makes the rather surprising statement that "there seems no virtue in trying to do a final and definitive study, or waiting until all the relevant evidence is in". Admittedly little detailed or systematic data was available to them at the time they published their views. Since Vanuatu Bislama predates by some time both Solomons Pijin and Tok Pisin, we will examine here some of the evidence provided by that variety of Melanesian Pidgin, after looking at the major claims of Mühlhäusler and Keesing and consider how this Vanuatu evidence can inform the debate.

For those not familiar with the key dates in the development of Melanesian pidgins, it is important to be aware that recruiting of labour for work in the plantations in Queensland, Samoa, Fiji and New Caledonia began in southern Vanuatu in 1863, in the Solomon Islands in 1872 and in what is now

Papua New Guinea in 1880. However, as early as 1847 Ben Boyd, who ran a whaling station at Twofold Bay in NSW, had recruited some 150 Melanesian labourers from Tanna (Vanuatu) and Lifou (Loyalty Islands, New Caledonia). Even at that time an early form of Bislama was spoken.

The *Australian* newspaper of 17 December 1847 describes a Tannese man in Sydney as "a quiet, unassuming, intelligent fellow, possessing a sufficient knowledge of broken English to make himself understood on a variety of occasions". He is quoted as saying that life in Sydney was not easy: "man Tanna no wife no nothing", whereas on Tanna there were "plenty bananas, plenty yams, plenty breadfruit, plenty sugar cane, plenty cocoa-nut and plenty pork – but in Sydney no nothing".

## 2. MÜHLHÄUSLER AND KEESING

The main areas of disagreement between Keesing and Mühlhäusler revolve around:

- (a) the status of Tok Pisin;
- (b) the role of substratum influence/calquing in the development of the various Melanesian pidgins;
- (c) the interpretation of features shared by languages such as Tok Pisin, Solomons Pijin and Bislama;
- (d) the time-scale of the expansion and stabilisation of what became Melanesian Pidgin;
- (e) the nature of the Pidgin pronominal system and the predicate marker.

Mühlhäusler has claimed a special and separate development for Tok Pisin as opposed to the other Melanesian pidgins (Bislama and Solomons Pijin). He attributes the distinctive features of Tok Pisin to the separate development of that language on the plantations of German Samoa in the 1860s and 1870s.

Keesing maintains that until the late 1880s the same dialect of Pidgin was spoken over a large region of the south-western and central Pacific, including the plantation areas of Queensland, Samoa, Fiji and New Caledonia and the recruiting areas (Vanuatu, the Solomons, and the Bismarks in Papua New Guinea). In other words he is denying Tok Pisin any special evolution prior to that time.

One of Keesing's major claims is that by the end of the 1880s, this regional Pidgin had acquired much of the syntactic complexity (and many of the specific lexical forms) characteristic of its daughter dialects in Papua New Guinea, the Solomons and Vanuatu (formerly the New Hebrides) in the twentieth century.

According to his reckoning (1988:112-113), the features common to all three daughter dialects of the late 1880s include not only a very substantial core of lexical items, but core syntactic patterns. These include:

- (a) the basic structure of the pronominal system including many patterns supposed to have developed in Tok Pisin since 1900; pronoun forms (if we view *olgeta* and *ol* as old variants of the same form) are essentially identical in all three dialects, and so is the structure of the paradigm itself, (including dual and plural sets, and a distinction between a first person non-singular inclusive and exclusive); so too are what have been taken to be specific historical anomalies in the Tok Pisin system, such as the use of the so-called 'predicate marker' *i* to maintain reference across clauses in discourse, specific constructional patterns such as *hem i* using this predicate marker, and the use of the third person plural pronoun (*ol* or *olgeta*) as a generalised plural marker;



- (b) the use of pronouns to embed relative clauses;
- (c) the systematic use of the transitive suffix *-im* to mark transitive verbs, and many of its specific uses (for example to create transitive verbs from statives);
- (d) the grammaticalisation of tense/aspect markers, such as the irrealis-marking *bai* (*babae*);
- (e) the use of periphrastic causatives using *mek-* and grammaticalised causatives using the transitive suffix *-im*;
- (f) a number of complementising and relativising constructions, using such forms as *see* and *wea* and *bulong*;
- (g) the use of *-fela/-pela* in a particular set of interconnected slots: suffixed to quantifiers (*tu-fela*, *sam-fela*); suffixed to a smallish class of attributive statives (*big-fela*, *gud-fela*); suffixed to demonstratives (*dis-fela*, *dat-fela*), used both as determiners and as pronouns; adverbially, apparently as a second, stative verb in compound-verbal constructions (*kilim strong-fela*); suffixed to pronouns or numerals, to indicate plurality of pronouns (*mifela*, *tufela*);
- (h) a whole set of clause-initial phrasal interrogatives (derived from English *what name*, *which way*, *what time* etc.);
- (i) a distinctive system of marking possession, using *bilong* to conjoin nouns, and nouns with pronouns to mark a genitive relationship metaphorically based on proximity;
- (j) what has been termed *ia bracketing* with relative clauses.

Keesing admits that a few of these features may represent parallel developments in Tok Pisin, Bislama and Solomons Pidgin, but that if this is the case, they could only have emerged out of a common dialect that already contained the essential materials. In fact the evidence adduced from Bislama will demonstrate the validity of this claim, although for reasons other than those which Keesing puts forward.

Keesing concedes, however, that while his hypothesis of a highly expanded Pidgin by the start of the 1890s is supported by what he terms "bits and pieces of textual evidence" attesting many of the features of subsequent Melanesian Pidgin, and by two substantial texts from the 1890s (Pionnier 1913, recording the Vanuatu Bislama of the 1890s, and Somerville 1893 for the Solomons), much of his case rests on distributional evidence. He states that where specific constructional patterns and forms are found in all three pidgin dialects, there are fairly strong grounds for their incorporation in the pidgin of the 1880s (since the New Guinea dialect was separated from the others in the latter years of that decade). As an example, Keesing inferred that on these distributional grounds *hem i* must have been present in the regional pidgin of the late 1880s. He reports (1990:3) that his claim was vindicated by the discovery of this sequence in a naval report by Commander E.G. Rason in 1895 after a visit to the Western Solomons. We will see later in the paper a kind of mirror image of Keesing's hypothesis, namely that the characteristics which distinguish each of the three Melanesian pidgins still exist in Bislama regional speech today, and that these same characteristics for the same areas of Vanuatu are recorded in the literature of the nineteenth century.

While Keesing maintains that until the 1880s there was spoken in island Melanesia an essentially unified Pacific Pidgin, characterised by a common grammar and lexicon, thus denying the separate genesis accorded to New Guinea Tok Pisin by Mühlhäusler, he argues that when the Solomons were separated from the New Hebrides (Vanuatu) and Queensland at the turn of the century, a fine tuning of the Solomons Pidgin pronominal system took place so as to match patterns in substratum

languages (having been generally based on an Oceanic substrate in the first place). Again, Vanuatu Bislama can inform this discussion (see below). He concludes, however, that the development of this pidgin is a phenomenon primarily of the nineteenth century, not of the twentieth (1990:15).

Mühlhäusler (1989a, 1989b)<sup>1</sup> disagrees with much of Keesing's hypothesis and conclusions. He states first of all (1989a:1) that Keesing misinterprets the nature of whaling, sandalwood and the beche-de-mer trade by arguing that they led to a "single early-Pidgin speech community" or "the linguistic community". Mühlhäusler maintains, rather, that there was very tenuous and often only indirect (via visiting Europeans) contact between the members of this postulated speech community. The nature of transmission of early Pacific Pidgin differed considerably from place to place and time to time: on board whaling and other vessels, adults learning from other adults in the plantations, children learning pidgin from returning adults as a second language, and in the early mixed beach communities. "That such different modes of transmission and crystallisation are signs of a single speech community or likely indices of a shared core grammar seems implausible" (1989a:1).

Mühlhäusler has more serious objections, however. He magnanimously states (1989a:2): "let us assume...with Keesing that the first occurrence of certain diagnostic constructions was followed by continuous diffusion and transmission". Of the ten constructions (listed above) which Keesing claims to have been common to South Western Pacific Pidgins in the late 1880s, Mühlhäusler maintains that the majority turns out not to have originated among speakers of Oceanic languages at all. He states that Philip Baker's provisional analysis of the corpus of data which he has collected suggests that the majority of them originated in Australia before Melanesian immigration to that country. He cites (1989a:2) the following:

1. The basic pronouns were first documented as follows:

<i>me</i>	I	New South Wales 1817
<i>yumi</i>	we incl	Queensland 1814
<i>yufela</i>	you pl	Queensland 1880
<i>alltogether</i>	they	Queensland 1858

2. The systematic use of the transitive suffix *-im* is first documented for New South Wales in 1826.
3. *-fela* as a suffix for quantifiers occurs in Queensland in 1848, with attributive adjectives in New South Wales in 1842 and with demonstratives in Queensland in 1842. Contrary to Keesing's assumption (1988:113), *-fela* was not introduced from China Coast Pidgin.
4. Phrasal interrogatives of the *whatname* type are first documented in Queensland in 1868.
5. The marking of possession by means of *bilong* first appears in New South Wales in 1826.

Mühlhäusler is not suggesting that there was a continuous transmission of these and other features. Rather, he is at pains to point out that such constructions could arise even where Oceanic substratum is absent. Given the quite considerable typological differences between Australian Aboriginal and Oceanic languages, the similarities of the Pidgin English used by their speakers will have to be explained in terms of linguistic universals or shared superstrate influence according to Mühlhäusler, rather than through the influence of substratum languages. The Australian materials presented by Mühlhäusler, however, run counter to his own earlier claim that most of these features

<sup>1</sup>Substantially two versions of the same article.

developed in Tok Pisin only in the twentieth century.<sup>2</sup> In the light of this, together with the evidence produced by Keesing, and the Bislama evidence discussed below, it would seem difficult to separate Tok Pisin from the mainstream developments in Melanesian Pidgins.

Another of the principal bones of contention for Mühlhäusler, though, is Keesing's claim that Pacific pidgins achieved a degree of stabilisation at a relatively early period. This is a matter of concern to Siegel also. Mühlhäusler finds that generally speaking, Keesing's approach suffers from a lack of distinction between chronological time and relative time. In Mühlhäusler's opinion, the former concept is needed to answer questions such as:

- (a) When did Tok Pisin become severed from the Melanesian Pidgin tradition (1880, 1890, 1900)?
- (b) When is a construction first documented for a particular area?
- (c) When were the first Pacific islanders repatriated from Queensland?

Separate from these issues is that of relative time. Mühlhäusler maintains that underlying his entire body of writings on the development of Pidgin English in the Pacific is the implicational or quantum-linguistic model which asks:

- (a) In what order do constructions (rule or rule environments) emerge in a pidgin language?
- (b) Does the presence of C imply that of B and A for a given lect?

Thus, with the third person plural pronoun, for instance, the question is not so much the chronological one of 'When is it first documented?' but rather 'If speakers use the third person plural pronoun will they also use the second person plural and the first person plural pronouns?'; 'If speakers use plural pronouns, will they also use dual pronouns'? Thus Mühlhäusler claims that he has found that for speakers of different ages in the same location, one could establish patterns such as:

D implies C implies B implies A

but that, at the same time, not all speakers have reached stages D and C. The importance of the implicational argument for the universals versus substratum debate is considerable. The claim is that such implicational patterns as the animacy or accessibility hierarchies, to use Mühlhäusler's words, provide principled limitations on what can be transferred from another language in what order (not, as Keesing (1988:171) interpreted it: "that substratum models will have an impact on a developing pidgin only at certain crucial points in its development"). It can thus provide a question that neither Keesing nor any other substratophile can answer: why is it that many constructions, rules or rule environments found in the substratum languages are not borrowed by pidgins, and why do those that are adopted get borrowed in a particular sequence rather than all at once?

The next problem which Mühlhäusler raises concerns continuity. He quotes (1989b:461) Hoenigswald (1971:476):

It is difficult enough to be quite sure, both in theory and in practice, when a given ordinary language is a descendant (under change) rather than a collateral relative of a given older language. It has been said that to discover a line of descent is to discriminate what has got handed down from mother to infant over the generations from what has passed through other channels. If this is true, the pidgins with their special mechanism of

<sup>2</sup>For this reason, perhaps, this Australian evidence does not appear in Mühlhäusler 1989b, a later version of the same review article.

exclusively secondary transmittal should indeed be troublesome to place on a family tree. And if it is further the case that pidgins are typically born and then again dropped from use in shortlived bursts of activity, the whole linear notion of 'gradual' change is not even a superficially useful approximation to the truth, as it is for normal, primary languages.

Thus the question of whether Pacific or Melanesian Pidgin is the same language either diachronically or geographically is one which has no easy answers.

As we have already seen, early in his book Keesing posits the existence of a Pacific pidgin which became fairly homogeneous and stable within less than a decade after 1845, that local dialects including Melanesian Pidgin began to emerge in the 1870s and that Tok Pisin became a separate dialect of Melanesian Pidgin around 1890. In addition he claims that most of the shared grammar of the various dialects of Melanesian Pidgin is of Eastern Oceanic origin.

Mühlhäusler (1989b:462) considers that there are three possible scenarios which could explain the position recognised by Keesing, as follows:

- (a) shared development up to 1880 with the additional stipulation that a number of crucial constructions were established firmly by 1880;
- (b) parallel development resulting from universal forces;
- (c) parallel development through continued borrowing from typologically similar Oceanic languages after 1880.

Mühlhäusler maintains that Keesing's line of reasoning often begs the question and that he indulges in circular arguments. He cites, for example, Keesing's statement:

*Him he gammon too much* was recorded in the Solomons in 1895. But the existence of these sequences in all the daughter dialects makes it almost certain that islanders were using such constructions by the 1880s. (1988:149)

Mühlhäusler warns of the dangers of using evidence from comparison of arbitrarily chosen synchronic states. In this context he demonstrates that some of what Keesing attributes to nineteenth century Melanesian Pidgin is lacking in Tok Pisin until the 1920s – the examples he chooses involve the use of *wininim* in comparative constructions and the sequences *mi mi* and *yu yu*. Mühlhäusler concludes that the kind of evidence Keesing produces can neither be taken seriously as a disconfirmation of the view that parallel development occurred as a result of universal factors nor as a proof of borrowing from substratum sources after 1880.

He goes on to attack the use of hypothetical data, especially with respect to pronouns, stating that what Keesing fails to present is hard and fast evidence that the inclusive/exclusive distinction with first person non-singular pronouns in fact existed at the time of his postulated crystallisation and splitting of Melanesian Pidgin. (It is in fact lacking from Pionnier's material of the early 1890s.) He argues further (1989b:467) that Borchardt's 'Guidance for Learning of Tok Boi' (1930) notes that the inclusive/exclusive distinction is "not commonly used" in the Tok Pisin of New Britain, although he lists it for plural pronouns. This he takes as evidence that the inclusive/exclusive distinction did not exist prior to 1890. (It is well and truly present in Jacomb 1914 for the New Hebrides, and in Fletcher 1923, reporting on the period 1912-1919).

Mühlhäusler also points out that the inclusive/exclusive distinction with non-singular first person pronouns is also lacking from Pidgin Fijian, where in fact it should have occurred as the speech community was composed almost entirely of speakers of Oceanic languages. He stresses also that

such a distinction was lacking from Samoan Plantation Pidgin as late as 1915, according to the most comprehensive source on this language. (I note however that the inclusive/exclusive distinction is lacking in Gilbertese, quite exceptionally, and that perhaps the pattern established in Samoa and some other areas was strongly influenced by Gilbertese speakers, who were known to have played an important part in the early formative years of Pacific Pidgin in Micronesia.)

Mühlhäusler raises a number of other objections to the evidence adduced and interpreted by Keesing, including the status of *-fela*. Do forms such as *me fello* or *him fello* (Keesing 1988:139) indeed stand for first person plural exclusive or third person plural as suggested by Keesing? Indeed the material and interpretations produced by both Mühlhäusler and Keesing are often ambiguous and cannot always be regarded as satisfactory.

Mühlhäusler concludes by challenging Keesing's earlier statements about the stabilisation of Melanesian Pidgin and the implication that by 1890 all Pacific Pidgins were at a relatively sophisticated level of development:

The amount of fluctuation and quite rudimentary pidgin documented (for Samoa and German New Guinea) for this period would seem strongly to contradict Keesing's suggestion. For instance many of the inhabitants of the Bismarck archipelago who worked in Queensland appear to have been unable to speak any form of pidgin and had to be interrogated via an interpreter by the 1885 Royal Commission. Other contemporary records, such as the Magisterial Records that I have examined, suggest a very limited knowledge of Pidgin English. (1989b:472)

Keesing (1990) replies to Mühlhäusler's criticisms, and those of a number of other reviewers, with a few noticeable shifts of ground from positions taken in his 1988 work, most notably in respect to claims of early stabilisation of Pacific Pidgin. (Siegel also takes Keesing to task for claiming that stabilisation had occurred by about 1860.) Keesing (1990:5) maintains that "stabilization (in the sense that I use the term) was a gradual process that was certainly still going on (in the New Hebrides, Solomons and New Guinea) in the early years of this century". I interpret this as an admission that we are still some way from being able to resolve many of the important questions concerning the development of Pacific/Melanesian pidgins, and that a *festina lente* approach might be more in order, rather than trying to have the last word as every tiny snippet of new information comes to hand. What is most striking is that the whole debate conducted by Mühlhäusler and Keesing, regardless of the stances taken, omits any but the most passing reference to Bislama, the pidgin variety spoken in Vanuatu, which was established some little time before either Solomons Pidgin or Tok Pisin. It is this situation which Tryon and Charpentier seek to rectify in their upcoming *Bislama and the development of Melanesian Pidgin*, a fine-grained study resulting from a data collection process which has been in progress for some 20 years, complemented by a fairly exhaustive search of archival documentary sources.

### 3. THE EVIDENCE FROM BISLAMA

When discussing Bislama and Bislama usage throughout Vanuatu, once one gets past the initial impression of homogeneity, largely because of the importance of Radio Vanuatu, one becomes increasingly aware of the quite considerable range of variation at all levels. Quite apart from the striking differences between the urban and rural registers of this lingua franca, one becomes aware of regional differences in phonology, lexicon and morpho-syntax. Indeed, there is so much variation



that once one focuses on the differences rather than the similarities in the varieties one encounters, one begins to realise that it is not really meaningful to talk about Bislama in general. One could better speak of Bislamas, such are the variations encountered. Thus the question 'Wanem Bislama'? We will now examine a few of the competing forms and structures from Vanuatu Bislama(s) and point up the important implications they have in the current debate about the evolution and development of Melanesian pidgins.

### 3.1 BISLAMA REGIONALISMS: MORPHO-SYNTAX

Nearly 30 Bislama morpho-syntactic regionalisms or departures from 'standard' Bislama are listed in summary form below. They have been observed particularly in the speech of older ni-Vanuatu (as citizens of Vanuatu choose to be called) in rural areas, although not necessarily limited to this category of speakers, since they are primarily regionalisms. The list is as follows:

	STANDARD	REGIONAL	MEANING	LOCALITY
1	<i>save</i>	<i>ken</i>	be able to know how to	NW Malekula
2	<i>ia</i>	<i>disfela</i>	this, that	Tanna, Ambrym, SE Pentecost
3	<i>wanem</i>	<i>want i</i>	desiderative, want, like	Epi, NW Malekula
4	<i>stap + VB</i>	<i>VB + stap</i>	habitual, continuative	Epi
5	∅	<i>bin</i>	narrative past tense	Efate, Erromango, Tanna, Aneityum
6	<i>olsem wanem</i>	<i>hao nao</i>	interrogative, how?	SW and NW Malekula
7	<i>ol i</i>	<i>olgeta i</i>	3 PL pronoun actor: they	SE Malekula
8	<i>wanem</i>	<i>laek</i>	desiderative, want to	Erromango, Maewo
9	<i>mi, yu</i>	<i>mi mi, yu yu</i>	1 SG and 2 SG actor: I, you	South Santo
10	<i>blong em</i>	<i>blong en</i>	3 SG possessive: his/her	Santo, Malekula
11	<i>luk</i>	<i>lukim</i>	see (Vtr)	Maewo, Malekula
12	<i>nomata</i>	<i>nating</i>	though, although	Malekula, Shepherd Is
13	<i>olsem wanem</i>	<i>wiswe</i>	interrogative, how, why?	Pentecost
14	<i>wea</i>	<i>weples</i>	interrogative, where?	Efate, Malekula
15	<i>wanem + N</i>	<i>eni + N</i>	whatever	Aneityum, Erromango
16	<i>yet</i>	<i>stil</i>	uncompleted action	Malekula
17	<i>we</i>	∅	relativiser	Erromango, Malekula
18	<i>ova long</i>	<i>bitim/winim</i>	comparative, more than	Efate, Paama
19	<i>long</i>	<i>wetem</i>	instrument, agent, with	Malekula



20	<i>ol</i>	<i>olgeta</i>	noun pluraliser	SE Malekula
21	<i>maet</i>	<i>mebi/ating</i>	dubitative	Tanna
22	<i>be</i>	<i>bat</i>	adversative, but	Erromango, Aneityum, Tanna
23	<i>-fela</i>	∅	adjective formant	Epi, Ambrym
24	<i>from we</i>	<i>bikos</i>	cause/reason, because, since	Efate/Shepherds
25	<i>N+gender</i>	<i>gender +N</i>	gender/sex, animals etc.	Malekula
26	<i>tastawe</i>	<i>from em ia</i>	consequence, therefore	Maewo, Epi
27	<i>kasem</i>	<i>tel</i>	until, as far as	Ambrym, Santo
28	<i>singsing</i>	<i>sing</i>	durative reduplication	Tanna
29	<i>talem long</i>	<i>talem</i>	verb + indirect object	Efate, Santo

The competing morpho-syntactic structures/functors listed above are indicative only and are far from representing the totality of morpho-syntactic variation in Vanuatu Bislama. In fact the table presented here is a considerable simplification since there are often more than two competing forms. They do however serve as an indication of the nature and extent of the variations encountered. At the same time, the distributional pattern of these morpho-syntactic competing forms is much more complex than the simple listing above suggests; for the geographical distributions of the morpho-syntactic isoglosses is not always clear-cut, since there are a number of overlaps. Yet again, there are often competing forms within a single geographical area, the variations being sometimes based on generational differences and sometimes seemingly genuine alternatives. A detailed survey is currently in progress, (Tryon and Charpentier, forthcoming) in terms of both synchronic and historical evidence for Bislama.

### 3.2 BISLAMA REGIONALISMS: LEXICON

There are also significant lexical differences for certain items in Bislama. While detailed isogloss maps are not yet completed, the following lexical differences between the Bislama of Santo and that of Malekula will serve to indicate the range of lexical regionalisms observed to this point:

ITEM	MALEKULA	SANTO
outside	<i>aotsaed</i>	<i>afsaed</i>
ocean	<i>blusi</i>	<i>dipsi</i>
community leader	<i>bigman</i>	<i>haeman</i>
tarpaulin	<i>tapolen</i>	<i>kanwis</i>
boar	<i>man pig</i>	<i>pig man</i>
material, cloth	<i>pis kaliko</i>	<i>haf kaliko</i>
nearly, almost	<i>klosap</i>	<i>samwe long</i>
coconut crab	<i>bigbol</i>	<i>krab kokonas</i>
until	<i>kasem</i>	<i>tel</i>

## 4. BISLAMA REGIONALISMS AND MELANESIAN PIDGINS

The regional variations exemplified above are indicative only of the range and scope of this phenomenon. They are the subject of an ongoing fine-grained study, the results of which will contribute to our understanding of the history and development of Bislama in Vanuatu. They can also inform the debate on the genesis of the other Melanesian pidgins, Tok Pisin and Solomons Pidgin, since many of the non-standard usages observed in Vanuatu Bislama today are diagnostic items in both of these languages, for example *ken* instead of standard Bislama *save* 'abilitative' and *laek* instead of *wantem* 'desiderative' (both characteristic of Tok Pisin); *hao nao* instead of *olsem wanem* 'how?', characteristic of Solomons Pidgin.

A number of the non-standard regionalisms listed above were at first considered to have resulted from contact with pidgin speakers from both the Solomons and Papua New Guinea, since there has been increasing communication throughout Island Melanesia. This hypothesis had to be rejected, however, when it was realised that the regionalisms in question characterise the Bislama of elderly speakers from remote rural communities who have rarely if ever travelled beyond their own area, and who would not have had the opportunity for contact with Melanesians from either the Solomons or Papua New Guinea. Nor have they had any exposure to English. More important, especially in terms of the debate about the genesis of Tok Pisin and Solomons Pidgin, is the fact that the selfsame regionalisms recorded today are recorded in the literature and archival materials of the nineteenth century for precisely the same areas of Vanuatu.

This remarkable regional conservatism in Vanuatu Bislama has significant consequences in terms of both Tok Pisin and Solomons Pidgin. The concordance of the archival materials of the last century and the morpho-syntactic and lexical regionalisms of Vanuatu Bislama today suggests that the other Melanesian pidgins, Tok Pisin and Solomons Pidgin, indeed shared a lengthy period of common development with Bislama. The evidence further suggests that they separated only after a considerable degree of stability had been achieved, differentiation and individual stabilisation being achieved after separation late in the nineteenth century.

## BIBLIOGRAPHY

- BICKERTON, D., 1975, *Dynamics of a Creole system*. Cambridge UP.  
 1981, *The roots of language*. Ann Arbor, Michigan: Karoma Press.
- CHARPENTIER, J.M., 1979, *Le Bislama(n) et le multilinguisme aux Nouvelles-Hébrides*. Paris: SELAF.
- CHARPENTIER, J.M. & D.T. TRYON, 1989, *Les problèmes linguistiques à Vanuatu*. *Ethnies* 8-9-10:13-17.
- CLARK, R., 1979, In search of Beach-la-Mar. *Te Reo* 22:3-64.
- CROWLEY, T., 1989, Referential and expressive expansion in Bislama. *English World-Wide* 10/1:85-118.  
 1990, *Beach-la-Mar to Bislama: the emergence of a national language in Vanuatu*. Oxford: Clarendon Press.
- DUTTON, T.E., 1985, *Hiri Motu: Iena Sivarai*. Port Moresby: University of Papua New Guinea Press.
- FLETCHER, R.J. ('Asterisk'), 1923, *Isles of Illusion: Letters from the South Seas*. Edited by Bohun Lynch. Constable, London.
- HOENIGSWALD, Henry M., 1971, Language history and creole studies. In Dell Hymes, ed. *Pidginization and creolization of languages*, 473-480. London: Cambridge University Press.
- JACOMB, E., 1914, *France and England in the New Hebrides: The Anglo-French Condominium*. George Robertson and Co. Melbourne.
- JOURDAN, C., 1985, *Sapos iumi mitim iumi: the social context of creolization in the Solomon Islands*. PhD thesis, The Australian National University, Canberra.

- KEESING, R.M.**, 1988, *Melanesian Pidgin and the Oceanic substrate*. Stanford: Stanford University Press.  
 1990, The expansion of Melanesian Pidgin: further early evidence from the Solomons. Mimeo.  
 in press, Reply to Belikov, Bickerton, Mühlhäusler, Romaine and Siegel. *Pacific Studies*.
- MOSEL, Ulrike**, 1980, *Tolai and Tok Pisin*. PL, B-73.
- MÜHLHÄUSLER, P.**, 1986, *Pidgin and Creole linguistics*. Oxford: Basil Blackwell.  
 1989a, Comments on Roger Keesing's *Melanesian Pidgin and the Oceanic substrate*. Mimeo.  
 1989b, Roger M. Keesing: *Melanesian Pidgin and the Oceanic substrate*, Stanford: Stanford University Press, 1988  
 [Review Article]. *Studies in Language* 13-2:459-475.
- PIONNIER, J.-N.**, 1913, Pigeon English ou Bichelamer, parlé universellement dans la Pacifique... *Revue de Linguistique et de Philologie Comparée* 46:109-117, 184-198.
- SIEGEL, J.**, 1987, *Language contact in a plantation environment: a sociolinguistic history of Fiji*. Cambridge: Cambridge University Press.
- SOMERVILLE, B.T.**, 1897, Songs and Specimens of the language of New Georgia, Solomon Islands (with an introduction by S.H. Ray). *Journal of the Royal Anthropological Institute* 26:436-453.
- TRYON, D.T.**, 1987, *Bislama: an introduction to the National Language of Vanuatu*. PL, D-72.  
 in press, Regionalisms and the history of Bislama. In R. Harlow and R. Clark, eds *VICAL 2*. Auckland: Linguistic Society of New Zealand.
- TRYON, D.T. & J-M. CHARPENTIER**, forthcoming, Bislama and the development of Melanesian Pidgin.

The first part of the report deals with the general situation in the country. It is noted that the economy is still in a state of depression, and that the government is facing a severe financial crisis. The report then discusses the various measures that have been taken to deal with the crisis, and the results of these measures. It is concluded that the situation is still very serious, and that further action is needed.

The second part of the report deals with the social situation in the country. It is noted that the standard of living is still very low, and that there is a high level of unemployment. The report then discusses the various measures that have been taken to deal with the social situation, and the results of these measures. It is concluded that the situation is still very serious, and that further action is needed.

The third part of the report deals with the political situation in the country. It is noted that the government is still in a state of instability, and that there is a high level of corruption. The report then discusses the various measures that have been taken to deal with the political situation, and the results of these measures. It is concluded that the situation is still very serious, and that further action is needed.

The fourth part of the report deals with the economic situation in the country. It is noted that the economy is still in a state of depression, and that the government is facing a severe financial crisis. The report then discusses the various measures that have been taken to deal with the economic situation, and the results of these measures. It is concluded that the situation is still very serious, and that further action is needed.

The fifth part of the report deals with the social situation in the country. It is noted that the standard of living is still very low, and that there is a high level of unemployment. The report then discusses the various measures that have been taken to deal with the social situation, and the results of these measures. It is concluded that the situation is still very serious, and that further action is needed.

The sixth part of the report deals with the political situation in the country. It is noted that the government is still in a state of instability, and that there is a high level of corruption. The report then discusses the various measures that have been taken to deal with the political situation, and the results of these measures. It is concluded that the situation is still very serious, and that further action is needed.

The seventh part of the report deals with the economic situation in the country. It is noted that the economy is still in a state of depression, and that the government is facing a severe financial crisis. The report then discusses the various measures that have been taken to deal with the economic situation, and the results of these measures. It is concluded that the situation is still very serious, and that further action is needed.

The eighth part of the report deals with the social situation in the country. It is noted that the standard of living is still very low, and that there is a high level of unemployment. The report then discusses the various measures that have been taken to deal with the social situation, and the results of these measures. It is concluded that the situation is still very serious, and that further action is needed.

The ninth part of the report deals with the political situation in the country. It is noted that the government is still in a state of instability, and that there is a high level of corruption. The report then discusses the various measures that have been taken to deal with the political situation, and the results of these measures. It is concluded that the situation is still very serious, and that further action is needed.

The tenth part of the report deals with the economic situation in the country. It is noted that the economy is still in a state of depression, and that the government is facing a severe financial crisis. The report then discusses the various measures that have been taken to deal with the economic situation, and the results of these measures. It is concluded that the situation is still very serious, and that further action is needed.

## QUESTIONS AND ANSWERS IN TOK PISIN

JOHN W.M. VERHAAR

### 1. PREAMBLE

Recent (and earlier) work on interrogativity has generated a measure of consensus among linguists about some of its properties which appear to be similar across languages. One of those is sequential position of interrogative markers: preverbal in VO languages and postverbal in verb-final languages (Lehmann 1973). Ultan (1978) shows that Subject-Predicate inversion as marking questions is not found in the majority of his sample (of 53 languages) and is even rare for yes/no (Y/N) questions; as for WH- questions, a majority of languages seem to favour clause-initial position of the WH-constituent, regardless of basic order type. In recent tradition, the distinction between Y/N (or 'polar') questions and WH- questions has become standard; for recent studies of seven languages, see Chisholm (1984).

Pidgin and Creole languages seem to share, across the globe, markers of questions not affecting word order, as well as a preference for phrasal WH- words which would be monolexical in most other languages (see Holm 1988:87); Tok Pisin (TP) illustrates this with *wanem samting?* 'what?', *long wanem hap?* 'where?', and other phrasal forms. Curiously, while inversion in TP questions occurs only for reasons unrelated to interrogativity (see example (58)), there is one case of inversion regularly occurring with *inap* 'can', 'be able to' (see example (12), below).

Most of the literature on interrogativity in TP is found in pedagogical grammars and reference works, and none of it goes into much detail. Some of the items are: Dutton (1973:40-43, 184-185, 231-233; 1985:2-4, 32, 50, 64-65, 236-237, 303-304); Laycock (1970: xxix-xxx); Mihalic (1971:15, 46); Mosel (1980:132-134); Mühlhäusler (1985:344); Woolford (1979:39-41, 43-48); and Wurm (1971:63-64). (On interrogative pronouns in older Samoan Plantation Pidgin, see Mühlhäusler (1978:96-97).)

The main part of this paper deals with interrogative clauses according to subdivision into Y/N questions and WH-questions. Section 2 deals with the former (including dependent Y/N questions) and with the answers to them; section 3 deals with the latter. Section 4 goes into WH-questions which have a serialised construction; section 5 discusses elliptical and parenthetical WH-expressions. Finally, section 6 deals with the grey area between interrogative and non-interrogative WH-constructions.

My data are in part taken from examples in the descriptive literature, in part from recordings I have made of student conversations, and also from a written corpus, consisting of a variety of self-help books and of the Bible translation. On reasons for including a translation in the corpus, see Verhaar (forthcoming a). In that paper, I have called the self-help books and the Bible translation together a 'virtual standard' (VS) of TP, for a variety of reasons, the most important ones being the remarkable homogeneity of this corpus grammatically, its consistent avoidance of colloquialisms, and its influence on literate Papua New Guineans. When this corpus is compared with my recordings (of geographical dialects: Madang Province, East Sepik Province, the Chimbu Highlands, and the Islands in the East), the difference between the two corpora in regard to the syntax of interrogativity is minimal (though in other regards – especially in regard to the use of the 'Predicate marker' *i* – there are significant differences).

I have given the data with full interlinear glossing, hoping thus to reach an additional readership not familiar with TP and not quite anglophone enough for easy analysis in all respects.

## 2. Y/N QUESTIONS

This section deals with: Y/N questions and their tags (2.1); answers to Y/N questions (2.2); and dependent Y/N questions (2.3).

### 2.1 Y/N QUESTIONS AND TAGS

Y/N questions in TP have the same word order as that of declarative sentences. Intonation will normally be enough to distinguish such sentences from declaratives; but the tag *o nogat?* (literally 'or not?') or *a?*, or *o?*, or *laka?*, may be added to Y/N questions (all end in falling intonation, except *o?*, which ends in level intonation). Both affirmative and negative Y/N questions may be introduced by *ating(?)* 'perhaps(?)', 'by any chance(?)', but it seems that *o nogat?* is only attached to affirmative sentences. The tag *a?* expresses surprise; *o?* expresses openness to a different answer (hence its characteristic level intonation at the end), and may be followed by *wanem* (Dutton 1985:64 see (10)). *laka* (Mihalic 1971:119; Dutton 1985:64; see (7) and (11), below) originates from the Islands and may now be heard on the mainland (but it doesn't occur in my written corpus nor in the recordings); it seems to confirm the quality of the sentence to which it is attached. Consider the following examples in which (1) is Highlands; (2) - (5), Madang; (6), Sepik:<sup>1</sup>

- (1) *Man ya em kisim pepa bilong em long UPNG (o nogat)?*  
 man DEM 3SG get title POSS 3SG at UPNG Q  
 Did this man (i.e. a lawyer) get his title at the University of Papua New Guinea?
- (2) *Bipo i gat dispela kain kastam tu (o nogat)?*  
 before PM be this kind custom too Q  
 Did you people formerly have that custom, too?

<sup>1</sup> Abbreviations used are as follows: ANT – anterior tense; AR – anaphoric reference; CL – classifier; DC – directional complement; DEM – demonstrative particle/pronoun; DUB – dubitative; EXCL – exclusive; FUT – future; HAB – habitual; INCL – inclusive; PFM – perfective marker; PM – predicate marker; POSS – possessive marker; Q – question marker.



- (3) *Em olsem wanem, yupela gat dispela kastam tu o?*  
 that as what 2PL have this custom too Q  
 But how's that, do you (people) have that custom too (or what)?
- (4) *Tru, a?*  
 true Q  
 True? (You don't say so!/I'll be darned!)
- (5) *Em tru, o?*  
 that true Q  
 Is that true (or what)?
- (6) *Em stret, a?*  
 that right Q  
 Is that (as I phrased it) right/correct?
- (7) *Ating em i samting bilong yu, laka?*  
 DUB 3SG PM thing POSS 2SG Q  
 Is that any of your business? (I certainly don't think so!)
- (8) *Ating Jon em i Krai onogat?*  
 DUB John 3SG PM Christ Q  
 Might John perhaps be the Messiah? (Luke 3:15)
- (9) *Yu no sem, a?*  
 2SG not be:ashamed Q  
 Aren't you ashamed of yourself?
- (10) *Yu nidim ka o wanem?*  
 2SG need car or what  
 Do you need a car or something else?
- (11) *Yu nidim ka a/laka?*  
 2SG need car Q/Q  
 You need a car, don't you.
- (12) *Inap yu tokim mipela wanem nem bilong tok ples bilong yupela?*  
 can 2SG tell 1PL.EXCL what name POSS language village POSS 2PL  
 Can you tell us the name of your local language?

Note that example (12) has inversion; it is found fairly frequently in my recordings, which do not seem to have any other examples of verb-initial position due to interrogativity.

## 2.2 ANSWERS TO Y/N QUESTIONS

Languages differ in their use of 'yes' and 'no' in reply to Y/N questions. I here summarise some points made in Verhaar (forthcoming b), which sets out various strategies for answers to Y/N questions.

Languages do not seem to differ in regard to the use of 'yes' and 'no' as answers to affirmative questions: 'yes' affirms the (affirmative) quality of such questions, and 'no' denies that quality. In reply to negative questions, however, let me distinguish the 'SAE' (to use Whorf's term for 'Standard Average European') strategy and the 'WPR' (Western Pacific Rim) strategy. In answer to

negative Y/N questions, the former uses 'yes' in favour of an affirmative answer, and 'no' for a negative answer; the latter, 'yes' for a negative answer, 'no' for an affirmative. The WPR strategy holds for many Austronesian languages, for Japanese, Chinese and TP. The name SAE should not be taken too seriously: many non-Western languages follow that strategy as well (for details, see Verhaar forthcoming b). Consider:

- (13) (a) *Em i kam pinis (onogat)?*  
3SG PM come PFM or not  
Has he come?
- (b) *Nogat.*  
no  
No(, he hasn't).
- (c) *Yes.*  
yes  
Yes(, he has).
- (14) (a) *Oi i no gat bikipela pe?*  
3PL PM not have large price  
They're not expensive?
- (b) *Yes.*  
no  
No(, they aren't).
- (c) *Nogat.*  
yes  
Yes(, they are).

Dutton (1985:236) reports that "many educated speakers" of TP now follow what is here called the SAE strategy in answering negative Y/N questions, that is, as illustrated in example (13). Given the continuing influence of English, this is not surprising. However, in my own data the WPR strategy is consistently followed, even in TP samples otherwise containing many anglicisms; also, speakers of my sample will (usually) answer such questions in the WPR mode even when speaking English.

### 2.3 DEPENDENT Y/N QUESTIONS

Dependent Y/N questions typically depend on a verb like *askim* 'to ask', and either follow that verb paratactically or are introduced by *sapos* 'if' – clearly a sample of grammatical anglicisation.

Consider the following examples (speakers from Madang Province):

- (15) *Mi laik askim yu bai yu mekim dispela samting o nogat.*  
1SG want ask 2SG FUT 2SG do this thing or not  
I want to ask you if you will do this (or not).
- (16) *Inap yu stori liklik long mipela sapos yu gat dispela kain*  
can 2SG tell a.bit to 1PL.EXCL if 2SG have this kind  
... *kastam olsem inisiesen*  
custom like initiation  
Can you tell us a bit if you people have this kind of custom like initiation?

Example (16) is the only case (in my recorded material) of *sapos*, which is normally conditional 'if', used to introduce a dependent question. (Anglicisation may not be the whole story, however, seeing that, across languages, conditionals and interrogatives are not far apart (Haiman 1978; Akatsuka 1985; Traugott 1985).) Speakers rejected the use of *sapos* in example (15), which may have something to do with the future tense/aspect of the dependent question.

### 3. WH-QUESTIONS

Independent WH-questions are discussed in section 3.1; dependent questions are often unstable as between interrogative and non-interrogative and are more conveniently treated as such, in section 6. (Elliptical and parenthetical questions are treated in section 5.)

#### 3.1 INDEPENDENT WH-QUESTIONS

##### 3.1.1 INVENTORY

TP WH-interrogatives are lexical, phrasal or clausal. I do not now consider as clausal the use of interrogatives elliptically, in such clauses as *husat?* 'who?' (with other information in context)' (on these, see section 5). It seems the only true clausal monolexical interrogative is *wasamara?* 'what (the heck) is the matter?' – a distinctly colloquial utterance (which apparently is becoming obsolete).

Lexically and phrasally, interrogatives are pronominal, adverbial or quantifying. Lexical pronominal interrogatives are *husat?* 'who?', *wanem?* 'what?'; the adverbial ones are, *we?* 'where', *hau?* 'how', *watpo?* 'why?', 'what for?'; and there is one quantifying interrogative, *hamas?* (dialectally also *hamaspela?*) 'how many' or 'how much?'.

*Husat?* is used substantivally only (except with *(ol) man*, and thus apparently only in Subject position; see also section 6); *wanem?*, either substantivally or adjectivally; in either case, sequential position is either clause initial or clause final, in part (but not exclusively) depending on what Argument position the interrogative takes.

As for phrasal pronominal interrogatives, I do not here count those that have Argument status (such as *wanem samting?*, *wanem kain?*, *husat (ol) man?*, or phrases marked by prepositions such as *long husat?* '(to / about / (etc.)) whom?', or *long wanem taim?* 'at what time?', 'when?') – those are essentially marked pronouns or pronominal phrases. Rather, considered here as 'phrasal' are the 'adverbial' ones: *olsem wanem?* 'how?', *bilong wanem?* 'why?'.

Of *we?*, *hau?* and *watpo?*, only *we?* may be clause final as well as clause initial.

*Husat?* is discussed in section 3.1.2; *wanem?* in section 3.1.3; *we?*, *hau?* and *watpo?* in section 3.1.4; phrasal WH-interrogatives in section 3.1.5.

##### 3.1.2 *husat?*

*Husat?*, as noted, is used substantivally only (except with *(ol) man*), is used semantically only [+Human], may be used as Subject, Predicate, Object, oblique Argument, as well as in all extranuclear positions, and it may take final position in all these slots (and must do so when used predicatively in equational clauses). *Husat?* may have plural reference and this may be made explicit in the use of *husat ol man?* 'who(PL)?'.

It is convenient to discuss equational clauses with predicative *husat?* separately (section 3.1.2.1), and other clauses with *husat?* after that (section 3.1.2.2).

### 3.1.2.1 PREDICATIVE *husat?*

In equational clauses, *husat?* in Predicate position has obligatorily final position. Equational clauses may be distinguished as 'identifying' (the '(This/it) is my father' type) and as 'characterising' (the 'He (anaphoric) is an artist' type). If *husat?* is the former, no Predicate marker *i* is used; if the latter, *i* is used in case the Subject is other than *mi* 'I', *yu* 'you (SG)', or *yumi* 'we (INCL)' (for details and discussion, see Verhaar, forthcoming a). Here, some discussion is needed for *Em* (*i*) *husat?*. Consider:<sup>2</sup>

- (17) *Yu husat?*  
2SG who  
Who are you?
- (18) *Yu husat man na yu save tok nogut?*  
2SG what.kind.of man *Sna* 2SG HAB talk evil  
Who are you to judge? (Romans 14:4)
- (19) *Dispela man em husat?*  
this man 3SG who  
Who is this man (whom I've never met yet)?
- (20) *Dispela man em i husat?*  
this man 3SG PM who  
Who is this man (whom I have met already)?
- (21) *Em husat?*  
DEM who  
Who is (it/that (person I've never met))?
- (22) *Em i husat?*  
3SG PM who  
Who is he (i.e. that person whom I've met already)?

(Note that *husat?* in example (18) is used adjectivally.) In examples (19) and (20), as well as in (21) and (22), the clause is identifying and characterising respectively; the gloss tries to approximate the difference (but in fact the questioner could well assume partial identification as present in the addressee).

The *husat?* data in my recordings reflect dialects (geographical and probably also social) largely doing without *i*, and thus do not seem to reflect (at least clause internally) any distinction between characterisation and identification. But biblical TP, as well as other texts of my VS TP corpus, reflects that distinction highly consistently (see Verhaar forthcoming a). I take some examples from that article here; I begin with a few examples with *mi*- and *yu*-Subjects just to illustrate that *i* is out in them (or present in case of constituents intervening between Subject and verb), without any relation to the distinction between characterisation and identification. Consider:

<sup>2</sup>*Sna* = serial *na*, on which see Verhaar, in press.

- (23) *Mi husat na mi inap pasim rot bilong God?*  
 1SG who *Sna* 1SG able lose way POSS God  
 Who am I to stop God? (Acts 11:17)
- (24) *Yu man, yu husat na yu bikmaus long God?*  
 2SG man 2SG who *Sna* 2SG answer.back to God  
 Who are you, man, to answer God back? (Romans 9:20)
- (25) *Dispela man em i husat na mi harim dispela tok long em?*  
 this man 3SG PM who *Sna* 1SG hear this talk about 3SG  
 [I had John [the Baptist] beheaded and] who [now] is this man [Jesus] [who some say is John] I hear these things about? (Luke 9:9)
- (26) *Na arapela man em husat?*  
 and other man 3SG who  
 And who is my neighbour? (Luke 10:29)
- (27) *Dispela Pikinini Bilong Man em husat?*  
 this son POSS man 3SG who?  
 This Son of Man, who is he? (John 9:26; 12:34)

In example (25), the context is represented in square brackets; the speaker is Herod, who has in part identified Jesus from all he has heard about him and now wants more information about him. In example (26), the speaker is the lawyer questioning Jesus about the command that one must love other people, one's 'neighbour': an identity question. In example (27), the John 9 text, the speaker is the man born blind, made to see by Jesus, whose identity as 'the Son of Man' he doesn't know; in the John 12 text, Jesus' hearers, having heard him say that 'the Son of Man must be lifted up', want to know who the Son of Man is. A glance at the Ramsey (1984) concordance, under the word *husat*, will turn up other examples in the New Testament and Psalms alone.

The rule illustrated here for equational *husat?* is copiously confirmed by non-interrogative examples, as shown in Verhaar (forthcoming a).

### 3.1.2.2 *husat?* IN OTHER CONSTRUCTIONS

Non-predicative slots taken by *husat?* are those of any Argument or non-nuclear constituent. Consider:

- (28) *Husat (ol) man i stap long dispela konperens?*  
 who PL people PM be at this conference  
 Who are attending the conference?
- (29) *Husat em i mama na brata bilong mi?*  
 who 3SG PM mother and brother POSS 1SG  
 Who is my mother and who are my brothers? (Mark 3:33)
- (30) *Husat i bin givim dispela namba long yu?*  
 who PM ANT give this authority to 2SG  
 Who has given you this right? (Mark 11:28)

- (31) *Yupela i laik mi lusim husat bilong dispela*  
 2PL PM want 1SG free who POSS this  
*tupela man i go long yupela?*  
 two man AR DC to 2SG  
 Which one of these two [Jesus or Barabbas] do you want me to set free for you?  
 (Matthew 27:21)
- (32) *Yupela i laik painim husat?*  
 2SG PM want look.for who  
 Whom are you looking for? (John 18:4)
- (33) *Profet i tok long husat na i mekim dispela tok?*  
 prophet PM speak about who Sna PM do this word  
 Whom is the prophet speaking about? (Acts 8:34)
- (34) *Em i pikinini bilong husat?*  
 3SG PM son POSS who  
 Whose son is he?

*Husat?* is Subject in examples (28) - (30); Object, in examples (31) and (32) (and then takes final position in keeping with basic VO order); and follows *long* or *bilong* in examples (33) and (34) respectively (then, too, final position is typical).

### 3.1.3 *wanem (samting/X)?*

As with *husat?*, it is also convenient with *wanem?* to treat equational clauses with predicative *wanem?* separately and after that clauses with *wanem?* in other syntactic positions (section 3.1.3.1). In both cases, 'absolute' *wanem?* is discussed along with *wanem samting?*. Finally, adjectival use of *wanem?* with a head other than *samting* (symbolised as X) is treated (section 3.1.3.2).

#### 3.1.3.1 PREDICATIVE *wanem (samting)?*

As with predicative *husat?*, so also with predicative *wanem (samting)?* a distinction is needed between identification and characterisation, and the use of the Predicate marker *i* has a similar function. But for *wanem (samting)?* there are three rather than two points on the continuum with identification and characterisation as extremes: the need for pure identification omits both *i* and *samting*; the need for what may be described as a mixture of further identification and initial characterisation has *i* but omits *samting*; finally, the need for characterisation only has both *i* and *samting*. The notions of 'identification' and 'characterisation' are rather global concepts, and perhaps a finer-grained continuum of 'topicality' would be better. But that requires a corpus of data in which questions figure prominently, and I do not have such data available in my corpus. Consider:

- (35) *Em wanem?*  
 DEM what  
 What (on earth) is that [thing I've never seen yet]?
- (36) *Em i wanem?*  
 DEM/3SG PM what  
 What is that/it? [i.e. that thing I seem to recognise but not quite]



- (37) *Em i wanem samting?*  
 3SG PM what  
 What is it? [i.e. I think I know, but need more information]
- (38) *Ombudsman em i wanem samting?*  
 ombudsman 3SG PM what  
 What is an ombudsman? [i.e. I think I know, but not quite]
- (39) *Biskit em i wanem samting?*  
 biscuit 3SG PM what  
 What is a biscuit? [i.e. I know more or less, but not quite]
- (40) *Man em i wanem samting na yu save tingting long em?*  
 man 3SG PM what            *Sna* 2SG HAB think of 3SG  
 What is man, that you should think of him? (Hebrews 2:6)
- (41) *Tok tru em i wanem samting?*  
 speech true 3SG PM what  
 What is truth? (John 18:38)

Note that *em* is demonstrative in example (35) and pronominal/anaphoric in example (37); in example (36) it could be either. The deicticity element is discussed in more detail in Verhaar (forthcoming a), and I will summarise the issues in a moment. Examples (35) - (39) are taken from Dutton (1973: 27), and what I have said about identification and characterisation rephrases substantially Dutton's explanation of the examples. For example (38), Dutton says that *samting* is added because *ombudsman* is a noun, but I believe the rule would hold also if the noun has occurred in previous discourse and is then picked up in pro-form.

The additional clarifications in square brackets are supposed to be the questioner's if he/she were to phrase them, but the matter is more subtle, involving what (the questioner thinks) the addressee knows. Thus, again, the matter is one of degree of topicality, in a stretch of discourse involving two (or more) speakers.

Predicative *wanem?* may also be headed by *kain* 'kind' (example (43) is from a recording (Madang TP):

- (42) *Em i wanem kain man na win wantaim wara*  
 3SG PM what kind man *Sna* wind together.with water  
*i harim tok bilong em?*  
 PM hear speech POSS 3SG  
 What kind of man is he, that even wind and water obey him? (Matthew 8:27)
- (43) *Na ol dispela spia ... em ol hap diwai o ain o wanem kain spia?*  
 and PL this spear 3SG PL CL wood or iron or what kind spear  
 And those spears...are they wooden spears, or iron, or what kind?

Here follows a summary of my claim set forth in Verhaar (forthcoming a). The 'Predicate marker' *i* is triggered by (wholly) 'non-deictic' Subjects, and thus *mi* and *yu* do not trigger it; neither does *yumi* 'we(INCL)', which combines [+1] and [+2], but *mipela* 'we(INCL)' does as it combines [+1] and [+3]. (*yupela* 'you(PL)' triggers *i*, because it is a *-pela* pronoun, and is thus 'nominal', in a sense explained in the paper, in addition to being pronominal.) Finally, *em* 'he/she/it' and (*em*) *ol*

'they' trigger *i* because [+3] is non-deictic – except when *em* is demonstrative, which it is in identifying (but not in characterising) equational clauses (for details, see the paper).

### 3.1.3.2 *wanem (samting/X)* IN OTHER CONSTRUCTIONS

As is the case with *husat?*, *wanem (samting/X)?* is found in all Argument as well as non-nuclear positions. Here are some data, example (47) is East Sepik dialect, examples without mention of source are from recordings or my notes:

- (44) *Wanem lo em i winim olgeta lo na i namba wan?*  
 which commandment 3SG PM exceed all commandment and PM number one  
 Which commandment is the most important and the first? (Mark 12:28)
- (45) *Wanem samting i namba wan? Gol ... o haus holi?*  
 which thing PM number one gold or house holy  
 Which is more important? Gold or the temple? (Matthew 23:17)
- (46) *Sapos ol i no mekim, bai wanem samting i kamap?*  
 if 3PL PM not do FUT what thing PM happen  
 If they don't do that, what will happen?
- (47) *Wanem samting i pasim rot na mi no inap kisim baptais?*  
 what PM close way *Sna* 1SG not can receive baptism  
 What is to keep me from being baptised? (Acts 8:36)
- (48) *Na si ol kisim ol save wokim wanem?*<sup>3</sup>  
 and sea.water 3PL get 3PL HAB make what  
 And the sea water they get, what do they do with it?
- (49) *Em long tok ples em yu kolim wanem?*  
 3SG in language village 3SG 2SG call what  
 What do you call that in the local language?
- (50) *Bai mi mekim wanem na mi kisim laip?*  
 FUT 1SG do what *Sna* 1SG get life  
 What should I do to obtain life? (Matthew 10:17)
- (51) *Yu tok wanem long yu yet?*  
 2SG speak what about 2SG self  
 What do you say about yourself? (John 1:22)
- (52) *Bai mi bihainim wanem rot?*  
 FUT 1SG follow what road  
 What road should I follow?
- (53) *Ol i gat oksen long wanem taim?*  
 3PL PM have auction at what time  
 When do they have the auction?

<sup>3</sup>*Si* is an anglicisation for *solwara* (non-anglicised *si* is 'surf' or 'high waves').

(54) *Dispela meri bilong wanem hap?*

this woman POSS what area  
What/which area is this woman from?

(55) *Bilong wanem samting tru yu wok long mekim?*

for what really 2SG DUR do  
What are you really doing that for?

From examples (45) - (47), it seems that *wanem?* in Subject position requires either *samting* or some other head constituent; that is, *wanem?* cannot be used 'absolutely' in this position. I have found no counter-evidence to this rule. In contrast, in Object position, *wanem?* can be used 'absolutely', as in examples (48) - (51); example (52) exemplifies attributive *wanem?* in Object position. Finally, examples (53) and (54) illustrate prepositional use, and *bilong wanem?* in example (55) is the normal phrase for 'why?'.

### 3.1.4 *we?*, *hau?* AND *watpo?*

These three non-pronominal WH-question words are quite unlike one another in that *hau?* 'how' is a recent anglicisation, and in that *watpo?* introduces an impatient question; only *we?* 'where?' is common and of high text frequency. Consider:

(56) *Watpo na mama i karim mi man?*

why Sna mother PM bear 1SG man  
Why did I have to be born a male?

(57) *Hau yupela save wokim mani long krokodail?*

how 2PL HAB make money from crocodile  
How do you make money from crocodiles?

(58) *Ol i stap we?*

3PL PM be where  
Where are they?

(59) *We stap papa bilong yu?*

where be father POSS 2SG  
Where is your father?

(60) *Yu stap we na yu kam?*

2SG be where Sna 2SG come  
Where are you coming from?

Example (56) is a jocular complaint by male boarding students, seeing that girls normally stay in the village. *hau?* in example (57) is from my recordings of students' conversations, but even in all of them it occurs only two or three times; the normal TP equivalent for 'how?' is *olsem wanem?*, on which see below (note the anglicised *krokodail* in example (57), rather than *pukpuk*; the dialect is Madang Province).

As appears from these examples, *we?* occurs clause initially and clause finally; in serialised constructions like example (60) it cannot be clause initial. As for the initial position, I have never found it for *we?* except with *stap*, a construction which is the interrogative variant of 'presentative

constructions'. The Predicate-Subject order is due to this type of construction, it seems, rather than to its being interrogative.

### 3.1.5 PHRASAL INTERROGATIVE ADJUNCTS

Phrasal interrogative adjuncts are: *bilong wanem (as / sam-ting) (tru)?*; and *olsem wanem?*. Examples are found in (3), above, and in examples (62) - (64) below. In addition, consider:

- (61) *Bilong wanem (as/samting) (tru) (na) yu mekim dispela samting?*  
 for what reason thing really Sna 2SG do this thing  
 Why (on earth) do you do this?

Observe that ('serial') *na* is optional here.

## 4. 'SERIALISED' WH-QUESTIONS

Several examples so far have had *na* glossed not as 'and' but as 'Sna', or 'serial *na*'. As proposed in Verhaar, (in press), 'serial' *na* in TP is different from *na* 'and', in that it functions either as a connective in serialisation or perhaps calques on clause chaining (found in many of TP's substrate languages) – a matter I leave open in that paper. Serialising *na* is not limited to WH-questions in this language, but it is to such questions that I limit myself in what follows. Examples of those in the present paper are: (18), (23) - (25), (33), (40), (42), (47), (50), (56), (60) and (61) (where *na* is optional). I add a few more; examples (62) - (64) are taken from Verhaar, (in press):

- (62) *Em i mekim wanem na yu gat sik?*  
 3SG PM do what Sna 2SG get sick  
 What did he do [to you] that you got sick?
- (63) *Olsem wanem na yu kam long mi?*  
 thus/as what Sna 2SG come to 1SG  
 Why do you come to me? (Matthew 3:14)
- (64) *Olsem wanem na em i no inap long wokim haus?*  
 thus/as what Sna 3SG PM not can to make house  
 Why can't he/she build a house?  
 How come he/she can't build a house?
- (65) *Bilong wanem na yupela i no laik go long pati?*  
 for what Sna 2PL PM not want go to party  
 Why don't you want to go to the party?

Example (62) parallels the construction in example (60), except that the first verb in example (62) is transitive. In examples (63) - (65), the first part is not verbal. This *na* after 'why?' words or phrases is so common in TP (though, interestingly, not in Bislama) that it reappears in local 'Englishes', in sentences like *Why and I trick you?* 'Why would I tell you lies?' (Yarupawa 1986: 50). For larger syntactic issues concerning 'serial' *na*, see Verhaar, (in press).

## 5. ELLIPTICAL AND PARENTHETICAL WH-EXPRESSIONS

Among WH-questions which are clauses, *wasamara?* 'what's the matter?' has already been mentioned; but that is a non-elliptical 'package loan' from English, which, also, cannot be broken down into constituent parts (it is an impatient question and, as noted, it is distinctly colloquial). Here follow some examples of frequently occurring elliptical questions (these may function as exclamations as well, and are as such perhaps parallel to rhetorical questions): *olsem wanem?* 'how did this happen?' / 'how are you?' / 'how are things?' / 'what's the matter?'; *bilong wanem?* 'why?' / 'what for?'; *long wanem hap?* 'in which area?' (or, simply, 'where?' – note that *we?* 'where?' cannot function elliptically); *bilong husat?* 'whose?' / 'for whom?'; *wanem?* 'what?'; *na wanem!* 'but of course!' is an exclamation.

An important use of *wanem* (but apparently rarely of *husat*) is one I may call 'parenthetical', in the form *wanem ya* (sometimes *wanem samting ya*) 'wotchamacolit' – it is found dozens of times in my recordings. It is, however, distinctly colloquial and I have not found it in my VS corpus. In my recordings, there is one example of parenthetical *husat ya* 'what's his name'.

## 6. DEPENDENT WH-CONSTRUCTIONS

In this final section, the heading is phrased as it is to avoid discussing dependent WH-questions only. The reason for this approach is obvious: across languages, dependent WH-questions overlap, in a fairly large 'grey area', with generalising WH-clauses (the 'whatever' / 'whichever' / 'wherever' type) and, in a number of languages, with relative clauses. Because of these overlaps, my treatment of these problems must be limited here. A few more generalist notes may precede discussion of data.

I am not aware of any wide-ranging studies on such 'grey areas' – let me designate them as 'grey WH-'. Within Germanic Romaine (1984) illustrates both the interrogative (WH-) and demonstrative (TH-) origin of relativisers in Germanic languages – the former being almost wholly and the latter being partly pronominal. I believe that (at least) the WH-origin of 'grey WH-' is much wider across languages, certainly in many Romance languages, but also Indonesian has fairly recently developed WH-relativisers (and 'grey' *siapa?*: *siapa?* → *siapa* (for [+Human]), while (*yang*) *mana?* → (*yang*) *mana* (for [-Human]/[+ Animate]) is older. In TP the *husat?* → (relative) *husat* change seems permanent, and anglicisation may not be the only source; curiously, 'grey' *husat* seems to be rare or nonexistent, and the 'greying' of *husat?* takes the form of the *wanem?* → (non-substantival for [+Human]) *wanem* change, as I will illustrate in a moment.

The type to be discussed here is the one of 'antecedent attraction': the type of relative clause illustrated in English by *Whatever books you buy, charge them to me*, in which the antecedent *books* is 'attracted' into the relative clause. Such relativisations are in the 'grey WH-' category presumably because they are close to constructions like 'What books do you want to buy? Charge them to me?'. The matter is discussed in detail in Verhaar (forthcoming a). Consider the following examples, (66) and (67) are from that paper, example (68) is from recordings, Sepik dialect:

(66) *Husat / wanem (ol) man i no laik harim dispela tok, ol i ken go.*  
 who-/what-ever PL man PM not want hear this talk 3PL PM may go  
 Whoever doesn't want to listen to the talk may go.

(67) *Wanem samting bai i kamap, tokim mi.*  
 whatever FUT PM happen tell me  
 Whatever happens, let me know.

- (68) *Ol bikman bai ol save long wanem samting ol save toktok*  
 3PL headman FUT 3PL know about whatever 3PL HAB talk  
 The headmen will then know what they are talking about.

Note that 'grey' *wanem* may be [+Human] (though not in VS TP). 'Grey WH-' relativisations are mandatory with non-third antecedents (*mi, yu*), which in this language cannot be relativised, thus resulting in 'whoever-among-you' constructions rather than in the 'you-whoever-you-are' type, (see Verhaar forthcoming a). Finally, there is also the 'adjunct' type of 'grey WH-', phrasally, in *long wanem* 'because', 'for which reason'.

## BIBLIOGRAPHY

- AKATSUKA, Noriko, 1985, Conditionals and the epistemic scale. *Language* 61/3: 625-639.
- CHISHOLM, William S., Jr, ed., 1984, *Interrogativity: a colloquium on the grammar, typology, and pragmatics of questions in seven diverse languages, Cleveland, Ohio, October 5th 1981 - May 3, 1982*. Amsterdam: John Benjamins.
- DUTTON, T.E., 1973, *Conversational New Guinea Pidgin*. PL, D-12.
- DUTTON, Tom [in collaboration with Dicks Thomas], 1985, *A new course in Tok Pisin (New Guinea Pidgin)*. PL, D-67.
- FISIAK, Jacek, ed., 1984, *Historical syntax*. Berlin: Mouton.
- GREENBERG, Joseph H., ed., 1978, *Universals of human language*. vol.4: *Syntax*. Stanford University Press.
- HAIMAN, John, 1978, Conditionals are topics. *Language* 54:31-89.
- HAIMAN, John, ed., 1985, *Iconicity in syntax*. Amsterdam: John Benjamins.
- HOLM, John A., 1988, *Pidgins and creoles*, vol.1: *Theory and structure*. Cambridge Language Surveys. Cambridge: Cambridge University Press.
- LAYCOCK, Don, 1970, *Materials in New Guinea Pidgin (Coastal and Lowlands)*. PL, D-5.
- LEHMANN, W.P., 1973, A structural principle of language and its implications. *Language* 49:47-66.
- MIHALIC, F., 1971, *The Jacaranda dictionary and grammar of Melanesian Pidgin*. Brisbane: Jacaranda Press.
- MOSEL, Ulrike, 1980, *Tolai and Tok Pisin: the influence of the sub-stratum on the development of New Guinea Pidgin*. PL, B-73.
- MÜHLHÄUSLER, Peter, 1978, Samoan Plantation Pidgin English and the origin of New Guinea Pidgin. *Papers in pidgin and creole linguistics* 1, 67-119. PL, A-54.
- 1985, Syntax of Tok Pisin. In Wurm and Mühlhäusler, eds 1985:341-421.
- RAMSEY, Evelyn M, 1984, *Konkodans, bilong painim tok insait long Buk Baibel*. Wewak: Christian Books Melanesia.
- ROMAINE, Suzanne, 1984, Towards a typology of relative-clause formation strategies in Germanic. In Fisiak, ed. 1984:437-470.
- TRAUGOTT, Elizabeth Closs, 1985, Conditional markers. In Haiman, ed. 1985:289-307.
- ULTAN, Russell, 1978, Some general characteristics of interrogative systems. In Greenberg, ed. 1985:211-248.
- VERHAAR, John W.M., in press, Serial *na* in Tok Pisin. (*Language and Linguistics in Melanesia*).  
 forthcoming a, The function of *i* in Tok Pisin.  
 forthcoming b, On *yes* and 'yes'.
- WOOLFORD, Ellen B., 1979, *Aspects of Tok Pisin grammar*. PL, B-66.
- WURM, S.A., 1971, *New Guinea Highlands Pidgin: course materials*. PL, D-3.
- WURM, S.A. and MÜHLHÄUSLER, P., eds, 1985, *Handbook of Tok Pisin (New Guinea Pidgin)*. PL, C-70.
- YARUPAWA, Shem, 1986, *Milne Bay informal variety of English*. Lae: Papua New Guinea University of Technology.



THE PROTO AUSTRONESIAN PHONEME \*t  
AND THE GROUPING OF THE AUSTRONESIAN LANGUAGES

JOHN U. WOLFF

1. INTRODUCTION

Professor George Grace's *The position of the Polynesian languages within the Austronesian (Malayo-Polynesian) language family* appeared during my days as a graduate student at Yale University. I chanced on this work in the linguistics seminar room one day and could not put it down until I had thoroughly studied it. Shortly thereafter I sought an interview with Professor Isidore Dyen and asked if I could work in Austronesian (AN)<sup>1</sup> linguistics for my thesis. I thought it would be fitting to present Professor Grace, who has done so much for Austronesian linguistics as a whole and whose work played a role in attracting me to Austronesian studies, with a study on Proto Austronesian phonology which, like Professor Grace's influential monograph, has important implications for the subgrouping of the Austronesian languages. In this paper I will discuss the evidence for the contrast between the purported phonemes \*C and \*t and what this means for the position of the Formosan languages in the Austronesian family.

Dahl (1973) proposed that all of the Austronesian languages with the exception of a few in Formosa formed a single subgroup as opposed to most (but not all) of the Austronesian languages in Formosa, and this view has found wide acceptance. In fact some scholars have concluded that there are three groups in Formosa and another group containing all the rest.<sup>2</sup> In any case the basic notion holds that the languages of the Philippines, which are located just south of the Formosan languages, are more closely related to those of faraway Hawaii and Easter Island, for example, than they are to their neighbours, Puyuma, Paiwan etc., to which they are very similar in grammar, vocabulary and even phonology. It is certainly not uncommon for languages to be more closely related to spatially

---

<sup>1</sup> Abbreviations used are as follows: At – Atayal, Bkl – Bikol, Bun – Bunun, Cb – Cebuano, Haw – Hawaiian, Ilk – Ilokano, Jav – Javanese, Kn – Kanakanavu, Kpp – Kapampangan, Mal – Malay, Mad – Madurese, Pai – Paiwan, PAN – Proto Austronesian, PF – Pre-Formosan, PFN – Proto Formosan, PRT – Proto Rukai-Tsou, PSF – Proto South Formosan, Puy – Puyuma, Rk – Maga Rukai, RkDd – Budai Rukai, Sai – Saisiat, Sam – Samoan, Sar – Saaroa, Tg – Tagalog, Tha – Thao, Ton – Tongan, Ts – Tsou.

<sup>2</sup> This is the view presented by Blust (1977), and almost everyone who has referred to the subgrouping of the Austronesian languages accepts this hypothesis, questionable as it is. Not, however, Tsuchida (1976:13), who considers the Formosan languages to form a subgroup within a group consisting of the western Austronesian languages (the 'Hesperonesian' group).

---

Robert Blust, ed. *Currents in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, 535-549. *Pacific Linguistics*, C-117, 1991.

© John U. Wolff

distant languages than to neighbours in the same group. French is very closely related to faraway Rumanian and not closely related to neighbouring Germanic languages. But in the case of French, German and Rumanian, French and Rumanian are remarkably similar in grammar, vocabulary and phonology and, despite more than a thousand years of contact, so different from German that there is little overt indication of a genetic relationship at all. This is the situation which often obtains: closely related languages tend to be similar and distantly related languages tend to be less similar, wherever they happen to be located. Certainly if the Philippine languages were more closely related to Malay, Polynesian, Moluccan etc. than they are to Puyuma and Paiwan it would be a unique situation, for languages like Kapampangan, Ilocano and even Tagalog are so similar in structure to Paiwan and Puyuma that one could almost translate from them into the latter on a morpheme-by-morpheme basis, and the portion of shared vocabulary in the basic items of the sort that tend to be old in a language (not just the 200-item Swadesh list, but a much broader vocabulary) is more like the portion shared by, say, French and Rumanian, than by French and German.

This is certainly enough to put under suspicion the theory of a Formosan higher order of proto-language. Many of the reasons advanced for this theory have been less than compelling. For the most part they have to do with items or processes which are lacking in the Formosan languages, but such facts tend to show that the Formosan languages form a subgroup, not that all the other Austronesian languages form a subgroup as opposed to them<sup>3</sup>. However, there have been two facts that have been totally compelling which, if true, could not be explained in any way other than that all of the languages outside of Formosa made important phonological innovations in which most of the languages of Formosa did not share. These are the phonemes *\*C* and *\*N*, for which most<sup>4</sup> of the languages of Formosa are thought to provide evidence, but which are thought to have merged with *\*t* and *\*n* respectively in all of the other Austronesian languages. If it should turn out (as I think it will) that the *\*C* and *\*t* sets are in complementary distribution and further that the *\*N* set is complementary to other sets, then the thread which holds the theory of a higher-order Formosan proto-language together is removed. No longer will one be able to defend a theory which implies a unique and difficult-to-believe situation.<sup>5</sup>

In this paper I will address only the *\*C* and *\*t* contrast. I believe that *\*N* is also non-contrastive, but that remains to be demonstrated. The distribution of the reflexes of *\*C* and *\*t* are complementary in terms of the accentual pattern of the root.<sup>6</sup>

<sup>3</sup>There is some morphological and syntactic evidence adduced, but I believe that in those cases it can be seriously questioned as to whether or not the fact adduced is an innovation or a retention. For example Blust (1977) adduces the second person singular genitive *-mu*, widespread in the extra-Formosan languages but unknown in Formosa, as an example of an innovation made in common by all of the languages outside of Formosa, but there is no compelling reason to say that the form *-mu* is innovated and not inherited from PAN.

<sup>4</sup>One must say 'most of' the languages of Formosa, not all of them, for Bunun, Amis, Kavalan, Siraya, Basay, Ketagalan (and Yami, a Batanic language) like the rest of the Austronesian family and unlike the other languages of Formosa, do not show a contrast between reflexes of *\*C* and *\*t*. Bunun, Kavalan, Basay and Ketagalan do not show a contrast between reflexes of *\*N* and *\*n*.

<sup>5</sup>Dahl (1981: section 13.3) thought that the various *\*H*, *\*Q* phonemes and the purported sibilant phonemes were further evidence for subgrouping the non-Formosan languages, but evidence for the existence of these phonemes is even weaker than for *\*N* and *\*C*. The contrasts which do seem to occur and are not explainable by borrowing or other sorts of contact phenomena are probably explainable in terms of prosodic features as our *\*C* and *\*t* contrast is.

<sup>6</sup>This notion is not original. I.I. Peiros of the Soviet Academy of Sciences suggested in a conversation with me that the accentuation combined with the phonological make-up of the root holds the key to unravelling the mystery of the *\*C* - *\*t* contrast. He was planning to work this out fully in his book *Problems of comparative-historical linguistics and linguistic history in east and south Asia (Problemy izucheniya sravitelinoistoricheskovo yazykoznaniya i lingvisticheskoi istorii vostoka i yuga Azii)*, hopefully to be published by the Soviet Academy of Sciences. This was

Of the Formosan languages we cite here, stress (or vowel length) is contrastive only in Kanakanavu and Rukai. Further, Rukai stress carries such a small functional load and has been subject to such strong analogical influence, that Rukai stress furnishes no evidence for stress in Pre-Formosan. Nevertheless, it is certain that the PF languages all had contrastive stress or length. Tsou, Rukai (Maga dialect) and Atayal retain traces of contrastive stress in the make-up of the root. In Atayal and Rukai if the penult was unstressed in the proto-language the resulting vowel in general is lost in the penult (phonetically becomes a schwa). If the penult was stressed, the resulting form has one of the vowels other than schwa. Atayal is not, however, a good witness for the stress pattern of the root in the proto-language for reasons discussed immediately below. Tsou has lost contrastive stress except on the morphophonemic level, that is, although phonetic stress is predictable in terms of the phonemic structure of the form in Tsou, the morphophonemic alternations involving vowel loss are motivated by morphophonemic stress position (Tsuchida 1976: section 2.3.1.3). Except for verbal roots, morphophonemic stress of the root in Tsou is taken to be evidence for the PF stress pattern.

There is a problem with demonstrating the hypothesis that the distribution of \*C and \*t is determined by the stress pattern of the root: the stress pattern in the PF root is not unequivocally determinable in terms of the present state of research on the Formosan and PAN stress. The stress patterns in all of the Austronesian languages are very heavily subject to analogical changes. Shift of stress is a morphological process in many of the languages, and compounding and affixation automatically shift the stress pattern of the word. Verb forms in Tsou and Kanakanavu rarely provide evidence for root stress because stress shift is part of the verbal morphology. In Tsou active verbs almost always show long-vowel roots, and suffixed verbs show shift in the stress from the first to the second syllable of the root. Further, compounding is so productive in Tsou, Kanakanavu and in Rukai that certain compounds have become part of the derivative morphology. Thus, a definitive determination of the stress pattern of PF forms will require much more detailed work on these languages than has been done heretofore. In short, there often is no agreement in stress patterns over languages. In fact, even within the same language and dialect alternative ways of stressing the root are common. Hence in many cases we can only make a hypothesis as to the stress pattern of the root in PF. There are, however, factors which enhance the probability that the hypothesis is correct. The following hypotheses come from detailed work with Cebuano and Tagalog and careful scrutiny of Tsou and Atayal data:

- (a) The stress patterns of nouns and some other forms which, other than stative adjectives, occur unaffixed tend to remain unchanged.<sup>7</sup>
- (b) In the Philippine languages in verbal roots the stress pattern of the actor focus verbs tends to reflect the inherited stress pattern. Verb forms in the Formosan languages rarely provide evidence.

---

not worked out completely in the version of the manuscript which he left with me, and the argumentation I adopt here is original. However, I would like to give Peiros credit for steering me to the notion that there might be a Verner's law-type solution to the \*C – \*t problem. David Zorc has also been working with this notion over the years.

<sup>7</sup>In Tsou morphophonemic stress in unaffixed forms of this sort is a good indication of the stress of the root in Pre-Formosan. For example, a form like *fatu* [fátu] 'stone' is pretty good evidence for a Pre-Formosan form with penultimate stress, \*bátu, and a form like *mcoo* [mcó] 'eye' is pretty good evidence for final stress in Pre-Formosan, \*maCá.

Thus, if there is agreement across languages in the stress pattern in these forms, the stress pattern can reliably be attributed to PF. A common situation is one in which there is disagreement among our Formosan languages, but the Philippine languages show agreement with some of them and confirm the hypothesis which we are seeking to demonstrate here. Such forms we will list here as confirming the hypothesis with the annotation that not all of the Formosan languages agree. However, we think the hypothesis is confirmed because the process of re-formation of accentual pattern due to compounding and affixation in the Formosan languages is overwhelming, and some disagreement in stress has to be expected.

### 1.1 ACCENT, STRESS AND LENGTH

The term accent in this paper refers to any stress or length phenomena. Stress is a combination of force, pitch, and length. We can reconstruct two types of roots for PAN: those which contained a long or stressed vowel on the penult<sup>8</sup> and those which did not. The actual phonetic facts are not reconstructable. Many of the current AN languages (e.g. Tagalog) distinguish long-vowel roots versus short-vowel roots. In some AN languages the contrast is between roots with a stress on the penult and those with no stress on the penult. We will talk about LONG-VOWEL ROOTS to refer to the type with a long vowel or stress in the penult and SHORT-VOWEL ROOTS to refer to the type with a short vowel or unstressed vowel in the penult. In any case PF maintained the distinction between these two kinds of roots. We hypothesise that the stress of PF functioned much as it does currently in the Philippines, where for words containing no long vowels (or no phonemic stress) the stress moves to the end of the word, and if there are suffixes or enclitics, there is no stress at all on the word and the stress is located on the enclitics; for example, Cebuano *ugát* 'veins' *ugatún* 'having prominent veins'. In short, in long-vowel roots the stress was on the long vowel, and in short-vowel roots the stress was at the end of the phonological word. We will transcribe long-vowel roots by writing an accent mark on the penult, and short-vowel roots by writing an accent mark on the final syllable.

### 1.2 THE RULE FOR THE DISTRIBUTION OF THE ALLOPHONES [\*C] AND [\*t] OF THE PHONEME \*t

The allophone \*C occurs in monosyllabic roots and in all short-vowel roots.<sup>9</sup> It also occurs in the initial and medial position of trisyllabic roots with a long vowel.<sup>10</sup> This rule applies irrespective of the position of the \*t phoneme within the word. However, for purposes of presenting the data, we will treat each position in the word separately. In other words, the rhythm of the root determined the distribution of the allophones of \*t, and further, \*t behaved differently in disyllabic roots than in trisyllabic roots. Disyllabic trochaic roots: that is, with a long penult (of the shape CÚ(C)CU(C), where U stands for vowel and C stands for consonant and parentheses stand for optional elements) had the allophone \*t. Disyllabic iambic roots: that is, with a short penult (of the shape CUCÚ(C) had

<sup>8</sup> Accent (contrastive stress or length) occurs only on the penult of the root in most AN languages which have retained this contrast. In most AN languages contrastive stress or length does not occur elsewhere in the root.

<sup>9</sup>Henceforth we will write \*C and \*t meaning the allophones [\*C] and [\*t] respectively of the phoneme \*t.

<sup>10</sup>As a matter of fact we have no examples of final \*t in long-vowel trisyllabic roots.

the allophone \*C. In trisyllabic roots the rule was the opposite: roots of the shape CUCÚ(C)CU(C) had \*C, whereas roots of the shape CUCUCÚ(C) had \*t.<sup>11</sup>

### 1.3 SOURCES AND DEFINITIONS

The materials here presented come from Tsuchida (1976). All forms which are cited in this work with PAN \*C, \*t or \*T are taken up here, with the exception of those which are confined to the Tsouic languages, for these can have no bearing on the hypothesis. I quote forms from Kanakanavu, Tsou, Atayal, Tagalog, Cebuano and from the Maga dialect of Rukai. Forms from other languages or dialects are only quoted if they are essential to the argument. Transcriptions for reconstructions, and for all Formosan forms except Atayal and Paiwan follow Tsuchida's transcriptions. Forms from Atayal and Paiwan and from the Philippine languages are transcribed as in the sources except that the glottal stop is transcribed with the symbol ʔ.

The reconstructions which we list are at the earliest level for which they can be posited. Not all are Proto Austronesian. Some are found only in Formosa, some in the Philippines and Formosa, some in the western languages, and some throughout the Austronesian area. This paper deals with a Formosan development, and it is not pertinent to the discussion how widely the forms are distributed. The reconstructions are transcribed as given by Tsuchida (except for immaterial typographic changes) with the following exceptions: we have omitted all subscripts, as they refer to facts not pertinent to our argument.<sup>12</sup> Also we do not transcribe \*H or \*W, as there is no good evidence that these proto-phonemes existed, and we transcribe \*T as \*t, and \*Q as \*q, as again there is no good evidence for a contrast between \*T and \*t, \*Q and \*q.<sup>13</sup> Also we transcribe \*D as \*d, for there is no contrast (Wolff 1975). The accent which we write on the reconstructed forms is the accent which we believe the form had in Pre-Formosan. We use the term Pre-Formosan (PF) to refer to the language or languages in which the two allophones of PAN \*t developed.<sup>14</sup>

## 2. DISYLLABIC SHORT-VOWEL ROOTS

Disyllabic short-vowel roots reflect the allophone \*C.

<sup>11</sup>In fact, we do not know that this was the stress pattern in trisyllabic roots rather than CÚCUCU(C). It does not matter. The main point is that only two accentual patterns can be reconstructed for PAN trisyllabic roots: one with a long vowel or stress on the penult of the root and one with no long vowel or no penultimate stress.

<sup>12</sup>However, Tsuchida's notations C<sub>1</sub> and t<sub>1</sub> refer to facts discussed in section 2.1.2.

<sup>13</sup>The tradition of hypothesising multiple proto-phonemes for correspondences which Dempwolff (1934-38) assigned to a single proto-phoneme stems from the clearly wrong-minded procedure of assigning a new symbol to every single irregularity in correspondence. Although Tsuchida (following Dyen 1971) states that these symbols are meant only as a notational device to indicate an irregular correspondence which needs to be explained in some way, in fact, he treats most of these symbols as a proto-phoneme, speaking of mergers and splits, and using these to define phonological environment. Other scholars have done the same. Doing so is tantamount to jumping to unjustified conclusions and forecloses the possibility of finding explanations for the irregularities which these correspondences symbolise in terms of analogy (including contamination and sound symbolism), borrowing or phonological environment. Surely only after the data from the languages showing the irregularity are completely understood could one give an explanation so definitive as to say that the irregularity is evidence for the existence of a new proto-phoneme.

<sup>14</sup>It is not germane to our argument whether PF was a subgroup or not. Because our hypothesis is that PAN \*t split into two phonemes \*C and \*t (this is a shared innovation), the PF languages might well be considered to form a subgroup. In fact, however, we are not forced to draw this conclusion. This was an innovation that may well have spread over language boundaries.



## 2.1 INITIAL POSITION

## 2.1.1 CASES IN WHICH THE ACCENT CAN BE RECONSTRUCTED UNEQUIVOCALLY

*CaNém	Sar <i>c-um-a-caḷəmə</i> 'bury'; Tg <i>taním</i> , Cb <i>tanúm</i> 'to plant'
*CaRéb/CaRúb (PSF)	Kn <i>c-um-a-carúvu</i> , Ts <i>m-a-crofə</i> 'cover oneself with a coverlet'
*CawíN <sup>15</sup>	Kn <i>cain-ána</i> , <sup>16</sup> Rk <i>cvéle</i> , Tg <i>taʔón</i> , <sup>17</sup> Ilk <i>tawén</i> 'year'
*Cebúng (PSF)	Kn <i>cuvúng-unu</i> , Ts <i>tro-cfung-a</i> 'confluence of rivers'; Pai <i>məcəvung</i> 'meet'
*CeRáb	Kn <i>c-um-a-cirávə</i> , Cb <i>tugʔab</i> , <sup>18</sup> Jav <i>atob</i> 'belch'; Tg <i>tigáb</i> 'gasp'
*Ciqáú	Kn <i>ciʔáú</i> , Ts <i>czou</i> 'kind of riverine fish'; Pai <i>ciqaw</i> 'fish'; Cb <i>tíʔaw</i> 'goatfish'
*Cumáy (PSF)	Kn <i>cumái</i> , Ts <i>cmoi</i> , Rk <i>cmée</i> 'a bear'

## 2.1.2 CASES WHICH REQUIRE DISCUSSION

For the following form there is no good evidence for the stress pattern of the root. (Because it is a verbal root, Tsou and Kananavu stress can provide no evidence.)

*Cungút (PSF)	Kn <i>apa-a-cúngunu</i> , Ts <i>cmunghu</i> , Rk <i>cunglu</i> 'join'
---------------	---

The following form probably had a short vowel, but there is no documentation. The Tsouic cognates are compounds, and Philippine languages show both short-vowel and long-vowel roots.

*Cawá	Kn <i>m-a-á-caca</i> , Ts <i>co-cvo</i> , Tg Cb <i>táwa</i> or <i>tawá</i> 'laugh'
-------	--

In the following case Tsou shows a long-vowel root, but Kananavu and the Philippines show a short-vowel pattern. Verbs in Tsou do not provide reliable information on the accentual pattern at an earlier stage.

*Capá	Kn <i>c-um-á-capa</i> , Ts <i>c-m-apo</i> 'roast over fire'; Tg <i>tapá</i> 'smoked meat', Cb <i>tapá</i> 'smoked fish'
-------	---

In the following case the Formosan forms reflect a short-vowel root, but the Philippine languages reflect a long-vowel root. In this case the PF roots can be taken to be with a short vowel, whatever the rhythm of the PAN root.

*Cubúq	Kn <i>cuvúʔu</i> 'bamboo shoot'; Tg <i>túboʔ</i> 'growth'
--------	---

<sup>15</sup>Tsuchida writes \*Cawil (PSF), not committing himself to the hypothesis that the forms outside of Formosa are cognate with the Formosan forms. However, I believe that the forms outside of Formosa (e.g. Mal *tahun*) can be connected with this form if we compare other PAN forms with the sequence \*awiC.

<sup>16</sup>In Kananavu if there is a stress on the first syllable of the suffix, this is taken to reflect a short-vowel root.

<sup>17</sup>The Tagalog forms require explanation. The ʔ is explained as follows: Tagalog sporadically reflects \*-w- and \*-y- with ʔ, e.g. *táʔo* 'person', *babáʔe* 'woman' etc. (cf. the Cebuano cognates which unambiguously reflect \*w and \*y: *táwu* 'person' and *babáyi* 'woman'). The vowel of the final syllable is explained by a hypothesis that PAN \*-win, which in Tagalog becomes -wun, becomes \*-wen in the Philippine languages (and perhaps in a wider group).

<sup>18</sup>The glottal stop in the Philippine languages may well reflect contamination from words having similar meanings. We do not reconstruct a PAN \*q in these forms because PAN \*q is reflected by a glottal stop in Kananavu.



In the following case the Philippine form has a short vowel in the penult of the root, but there is no evidence from the Formosan languages, and only the Philippine form provides evidence for the PF accentuation.

\**Cazém*                      Sai *s-əm-aəm*, RkBd *cáiti*, Tg *talím* 'sharp'

In the following case the only evidence for the root accentuation is from the Philippine languages, and the accent pattern in the Philippines is not unanimous.<sup>19</sup>

\**CuNú*                      Kn *c-um-ú-cunu*, Ilk *tunoén* 'roast, broil'

In the following case Tsouic and Atayal show reflexes of a long vowel in the penult, whereas Rukai shows a reflex of a short vowel. We believe that Rukai reflects the original PF stress pattern and the other languages developed a long vowel on the penult.

\**Cábu* (PFN)                Kn *cavu-cavúa* 'rice cake'; Ts *c-m-ofu*, Rk *o-cbóo*, At *smabo* 'wrap'

In the following cases Tsou, Kananavu, and some of the Rukai dialects show reflexes of the \**t* allophone, whereas the other Rukai dialects and other Formosan languages which have separate reflexes for \**C* show the reflexes for \**C*. Our theory allows us to explain these forms as having developed different stress patterns, that is, the forms which show the \**t* allophone developed a long vowel in the penult: the Maga dialect of Rukai which retains evidence of the stress pattern of the root retains the vowel of the penult, indicating a long vowel in the penult.<sup>20</sup>

\**CalíS*                      Kn *talísi*, Ts *tresi*, RkBd *cálisi*, Tg Cb *táli?*<sup>21</sup> 'rope'

\**Cangís*                      Kn *tumátangi*, Ts *mongsi* (< \**tmongsi*), Pai *cmangit*, Tg *tángis*, Cb *tangís* 'cry, wail'

\**Caqí*<sup>22</sup>                      Kn *táaʔi*, Ts *tʔee*, Pai *caqi*, Cb *táʔi* 'faeces'

\**CaSíq*                      Kn *tumataʔisi*, Ts *tmeʔsi*, Pai *cmagis*, Tg *tahíʔ* 'sew' (Kn and Ts reflect metathesis of \**S* and \**q*.)

\**CebúS*<sup>23</sup>                      Kn *təvəsə*, Ts *təfsə*, RkBd *cubúsu*, Tg *tubó*, Cb *tubú* 'sugarcane'

\**Cingás*                      Ts *ru-ngtosə*, RkBd *mua-cíngasə*, Rk *mu-tingásə* 'remove food particles from between the teeth'; Pai *cingas*, Cb *tingá* 'food caught in teeth'

<sup>19</sup>It would be worth investigating a theory that an accent on the prefix of the Kananavu form, as is the case here, is an indication of a short-vowel penult.

<sup>20</sup>Dahl (1981:21) suggests that the reflexes within these cases could be explained by a rule of dissimilation in roots with sibilants. The problem with this explanation is that these irregularities are not confined to roots with sibilants, and the forms with /t/ are not all found in the same language. Still, it would be worth investigating exhaustively the phonology of the languages which show this irregularity, as there seem to be some restrictions on the occurrence of affricates and sibilant /s/ in roots in these languages.

<sup>21</sup>The final glottal stop in the Philippine reflexes is unexplained.

<sup>22</sup>Kananavu and Cebuano reflect \**táqi*, and this is a possible reconstruction. Paiwan reflects \**Caqí*, and Tsou *tʔee* reflects a short-vowel penult, but the initial /t/ is anomalous.

<sup>23</sup>Tsuchida reconstructs \**tebuS*.

## 2.2 MEDIAL POSITION

## 2.2.1 CASES IN WHICH THE ACCENT CAN BE RECONSTRUCTED UNEQUIVOCALLY

* <i>maCá</i>	Ts <i>mcoo</i> , Rk <i>mcáa</i> , Tg Cb <i>matá</i> 'eye'
* <i>maCéy</i>	Kn <i>maa-macái</i> , Ts <i>mcoi</i> , Tg Cb <i>matáy</i> 'die'
* <i>NaCéng</i>	Kn <i>naténgə</i> , Rk <i>lcéngə</i> , Ilk <i>naténg</i> 'vegetables'; Bkl <i>natóng</i> 'taro leaves'
* <i>paCéy</i>	Kn <i>m-ia-pacái</i> , Ts <i>o-pcoi</i> , Tg Cb <i>patáy</i> 'kill'; Rk <i>ma-pcéé</i> 'wither'
* <i>puCáq</i> (PSF)	Kn <i>pucáʔə</i> , Pai <i>pucaq</i> 'bubble, foam'
* <i>qaCáng</i> (PSF)	Kn <i>ʔacángə</i> 'stone walls'; Pai <i>qacang</i> 'pigpen'
* <i>qaCéy</i>	Pai <i>qacay</i> , Tg Cb <i>atáy</i> 'liver'
* <i>qaCí</i>	Kn <i>ʔum-á-ʔaci</i> , Rk <i>u-cfi</i> 'dam up side stream to catch fish'; Bkl <i>atí</i> 'low tide'

## 2.2.2 CASES WHICH REQUIRE DISCUSSION

In the following case Atayal reflects a long-vowel root, but all the other forms reflect a short-vowel root. There may well have been alternative accentuation of this root in PF, and this accounts for the reflex /t/ in Kananavu, Rukai, Tsou and Atayal.

* <i>paCás</i> <sup>24</sup> (PFN)	Kn <i>tapásə</i> , Ts <i>ta-tpos-a</i> 'pattern, design'; Ts <i>tposə</i> 'letter, book'; RkBd <i>ua-pacásə</i> 'write'; Rk <i>ptásə</i> 'embroider'; At <i>patas</i> 'tattoo'
------------------------------------	--

In the following case Tsou and Rukai reflect a short-vowel root, but Kananavu reflects a long vowel. The Philippine cognates have long-vowel penults. It is nevertheless likely that Tsou and Rukai forms reflect the PF accentuation.

* <i>kuCú</i>	Kn <i>kúucu</i> , Ts <i>ʔcuu</i> , Rk <i>kcúu</i> , Tg <i>kúto</i> , Cb <i>kútu</i> 'head louse'
---------------	--

In the following case Cebuano has a long-vowel root, but Tsou reflects a short-vowel root. The Bikol accent is not given in the sources.

* <i>qaCéb</i>	Ts <i>cəfə</i> , Cb <i>átub</i> , Bkl <i>atob</i> 'deadfall trap'
----------------	---

In the following case the stress pattern cannot be reconstructed. There has been contraction of the root with the adjective-forming prefix *\*ma-* in many of the forms. Another factor which affects the outcome is that the *\*t* may have been preceded by /n/ in PF. Since in PF closed syllables were treated like long-vowel syllables, a *\*t* preceded by *\*n* would be reflected as /t/.

* <i>qetaq/qentaq</i> <sup>25</sup>	Kn <i>matáʔə</i> , Ts <i>mato</i> , Sai <i>mantäʔ</i> , Bun <i>matʔah</i> , <sup>26</sup> Mal <i>mentah</i> , Ton <i>mata</i> 'unripe, uncooked, raw'; At <i>qmataq</i> , Ton <i>ʔota</i> 'eat raw'
-------------------------------------	---

No stress pattern can be reconstructed for the following case either.

* <i>CekeS</i> (PSF)	Pai <i>cəkəs</i> 'kind of bamboo'
----------------------	-----------------------------------

<sup>24</sup>Or we could reconstruct *\*pátas*; Tsuchida reconstructs *\*patas*.

<sup>25</sup>Tsuchida reconstructs *\*ma(n)taq*.

<sup>26</sup>Quoted by Dahl (1981:146).

### 2.3 FINAL POSITION

#### 2.3.1 CASES IN WHICH THE ACCENT CAN BE RECONSTRUCTED UNEQUIVOCALLY

- \**kułíC* Kn *kuíci*, Ts *rici*, Tg *kulít* 'peelings'  
 \**LekéC/zekéC* Kn *ma-ta-nəkəcə*, Ts *hʔərc-a* 'sticky'; Pai *sə-djəkəc* 'adhere'; Tg *dikít*, Cb *dukút* 'stuck, adhering'  
 \**SeRéC*<sup>27</sup> Kn *ma-ʔərcə*, Ts *s-m-o-rəcə* 'tie tightly'; Rk *u-rəcə* 'tie two strings'; Mal *herat* 'tiød tight'

#### 2.3.2 CASES WHICH REQUIRE DISCUSSION

In the following cases Tsou does not agree with the Philippine stress, but the Philippine languages provide the most reliable evidence for reconstructing stress patterns in the first example because it is a verb. The second example is a noun and we take the Tsou form to provide the evidence for the PF stress pattern.

- \**kaRáC* Sar *um-a-aracə*, Ts *b-orcə*, At *kmat*, Tg Cb *kagát* 'bite'  
 \**qañíC* Sar *ʔaíci*, Ts *hici*, Cb *ánit* 'leather'

In the following case Rukai retains the vowel of the first syllable. The loss of the vowel of the first syllable perhaps does not apply to /u/.

- \**uRáC* Kn *urácə*, Ts *vrocə* 'blood vessel, vein'; Rk *uvácə*, Tg Cb *ugát* 'artery, vein'

In the following case the accent in PF cannot be reconstructed.

- \**rengəC* (PSF) Kn *ara-ca-rəngəc-a*, Pai *ma-rəngəc* 'jealous'

#### 2.3.3 COUNTERCASES

In the following case the Philippine languages have a long-vowel root.<sup>28</sup>

- \**langíC*<sup>29</sup> Puy *ringaT*, Tg Cb *lángit*, Mad *langngeʔ* 'sky'

### 3. DISYLLABIC LONG-VOWEL ROOTS

Disyllabic long-vowel roots reflect the allophone \*t.

<sup>27</sup>Tsuchida reconstructs \**qeReC*, as he did not have the Philippine cognates. The correspondence Philippine /h-/ , Kananavu /ʔ-/ is not explained. The Kananavu initial glottal stop developed in word-initial position before a vowel. There are several forms in the correspondence Philippine initial /h/ – Tsouic nothing to which Dyen and Tsuchida assigned the symbol S<sub>2</sub>.

<sup>28</sup>Zorc (1978) believed that Madurese *langngeʔ* reflects \**langít* (because of the doubled medial consonant). However, there is no strong correlation between gemination in Madurese and root final stress in other languages.

<sup>29</sup>This form is not quoted by Tsuchida.

## 3.1 INITIAL POSITION

## 3.1.1 CASES IN WHICH THE ACCENT CAN BE RECONSTRUCTED UNEQUIVOCALLY

* <i>tápis</i>	Sar <i>tapisi</i> 'man's skirt'; Tg Cb <i>tápis</i> 'cloth wrapped around waist worn like a skirt'
* <i>táRa</i>	Kn <i>t-um-á-tara</i> 'ambush'; Ts <i>m-oo-tro</i> , At <i>m-naga</i> 'wait'; Ilk <i>tága</i> 'await'
* <i>télu</i>	Kn <i>u-túlu</i> , Ts <i>туру</i> , Rk <i>túru</i> , Cb <i>tulu</i> <sup>30</sup> 'three'
* <i>témuy</i> (PFN)	Kn <i>?i-túmulu</i> 'many, much'; Ts <i>o-tmuzu</i> 'eat much'; <i>pe-tmuzu</i> 'drink much'; Puy <i>ma-təmuy</i> 'full'
* <i>téNuq</i>	Kn <i>təná?e</i> 'resin, sap'; Rk <i>túlu</i> , Cb <i>tínu?</i> or <i>tunú?</i> 'coconut milk' <sup>31</sup>
* <i>túduq</i>	Kn <i>tucú?u</i> , RkBd <i>ua-túDu</i> , Tg <i>túlo?</i> , Cb <i>túlu?</i> 'leak'

The following forms have a suffix. If the suffixed root has a long penult, we hypothesise that in PF the root alone had a long vowel.<sup>32</sup>

* <i>taRúqan</i> (root * <i>táRuq</i> )	Kn <i>tarú?anə</i> , Ts <i>trova</i> 'shelter used when hunting'; Rk <i>tovnáa</i> 'shelter in fields'; Tg <i>tágoq</i> 'hide'; Cb <i>táguq</i> 'put, hide'
---	---

Roots with a medial consonant cluster were treated like long-vowel roots and reflect the allophone \**t*.

* <i>taktak</i>	Kn <i>t-um-a-takətakə</i> 'hew down'; Ts <i>m-o?to?ə</i> 'clear land for swidden'; Kpp <i>taktak</i> 'cut grass with bolo'
* <i>téRbeS/tíRbeS</i> (PSF)	Kn <i>tərvəvəsə</i> , Ts <i>trəfsə</i> , Rk <i>tíbsə</i> <i>Zelkova formosana</i>

The resemblance of the following forms may have to do with sound symbolism rather than inheritance.

* <i>tuktuk</i>	Kn <i>m-akí-tuku</i> , Ts <i>m-?o-t?u-t?u</i> , Rk <i>u-tkútku</i> 'pound'; Jav <i>TuTuk</i> 'knock'
-----------------	--

The following form should be reconstructed as a disyllabic root consisting of a doubled syllable.<sup>33</sup> The \**t* allophone occurs because the root had a medial consonant cluster.

* <i>witwit</i>	Kn <i>m-ari-itítí</i> 'wave'; Ts <i>reu-vtívtí</i> 'move and swing tail or ears (of animal)'; Tg <i>witwit</i> 'shake forefinger in scolding'; Bkl <i>witwít</i> 'mechanical swing ride in carnival'
-----------------	--

<sup>30</sup>In the Philippine languages roots which had \**e* in the penult in PAN always stress the final syllable.

<sup>31</sup>In Cebuano in the case of nouns which refer to the result of the action of a verb and which occur with both penultimate and final stress, it is probable that the form with the penultimate stress is original, as there is an analogical tendency to give final stress to nouns of this class.

<sup>32</sup>This is the process which is productive in the Philippine languages. The accent stays on the penult of the affixed form if the form without the suffix has penultimate accent, e.g., Tagalog *pútol* 'cut', *putúlan* 'cut from'. If the accent is on the final syllable of the root, the accent of the suffixed form is also on the final syllable, e.g. Tagalog *inóm* 'drink', *inumín* 'drink it'.

<sup>33</sup>Tsuchida reconstructs \**witwit*, but the middle vowel is epenthetic. In the Philippine languages doubled monosyllables sporadically develop stressed epenthetic vowels, where the epenthetic vowel is /i/ if the root vowel is /i/ and otherwise /u/. In this example Bikol shows a reflex with the epenthetic vowel, C<sub>1</sub>UC<sub>2</sub>ÚC<sub>1</sub>UC<sub>2</sub>, whereas Tagalog shows a reflex without the epenthetic vowel, C<sub>1</sub>UC<sub>2</sub>C<sub>1</sub>UC<sub>2</sub>.

### 3.1.2 CASES WHICH REQUIRE DISCUSSION

In the following case the Formosan forms indicate a long-vowel root but the Philippine languages reflect a short-vowel root. As in the case of \*Cubúq in section 2.1.2 above, the PF stress pattern is reflected by the Formosan forms, whatever the PAN stress pattern was.

\*taSáNan (root táSaN) Kn *tanásə*, Ts *hos-a* 'village'; Tg *tahán* 'reside'

In the following cases there is no good evidence for the stress in PF.

\*tákid Kn *m-aa-takíci*, Kpp *tákid*<sup>34</sup> 'adhere'

\*táLam Kn *ku-a-tanámə*, Ts *oo-thomə*, At *tmalam* 'taste'

\*túqaS Tha *tuqa-tú:qaš*<sup>35</sup> 'old (of age)'; Ton *motuʔa* 'old man'

In the following example Kanakanavu reflects a long-vowel root but the Philippines a short-vowel root.

\*túqed Kn *t-an-əʔəcə*, Tg *tuʔód*, Cb *tuʔúd* 'stump of tree'

### 3.1.3 COUNTERCASES

In the case of roots with open penult there is only one counterexample and three other forms which look like counterexamples, but are not. First the counterexample. This form may be a counterexample, but it is not a strong one. To be sure, the Philippine languages show a short-vowel root, but in this case the unaffixed root alone in the Philippine languages refers to the result of the action of the verb, a type of noun which normally has a short vowel in the penult. This short-vowel pattern is often generalised throughout the paradigm, and roots of this type which show a short-vowel penult may well have had a long vowel at an earlier stage. Rukai also points to a short vowel in the penult, and the explanation for the Rukai phenomenon may be the same.

\*tapÉS Kn *t-um-a-tapəsə*, Ts *m-opsu*, Rk *o-tpəsə*, Tg *tahíp*, Cb *tahúp* 'chaff'

The following cases look like counterexamples, but definitely are not so. In the first example, the forms compared are not cognate: their resemblance is coincidental.

\*tu(m)bák Kn *tuvákə*, Rk *tbákə* 'cowrie'; Haw *kupa* 'kind of red cowrie'<sup>36</sup>

The second example consists of a doubled monosyllable, but the forms cited may not be related.

\*CengCeng Rk *i-cngəcngə* 'boy or girl of right age to be married'; Pai *cəngcəng* 'enough, fitting, just right'; Cb *tungtung* 'reach a certain step, age, degree'

The third example includes forms which have the \*t reflex in Saaroa and Rukai but the \*C reflex in Pawai. In any event, this is a form which is likely to be affected by sound symbolism and does not provide evidence for or against our hypothesis.

\*CugCug (PSF) Sar *m-utu-tukutuku*, Rk *ruu-tgútgu*, Pai *ma-cugcug* 'bumped on head'

<sup>34</sup>The sources do not provide the accent for this root. Blust (pers. comm.) states that the Kapampangan form is *tákid*, which supports my reconstruction.

<sup>35</sup>The vast majority of the Thao forms quoted by Tsuchida have penultimate stress.

<sup>36</sup>Biggs (1979) reconstructs Proto Polynesian \**tupa* 'land crab' and the Hawaiian form refers to a dark red cowrie. The resemblance to the Formosan forms is purely coincidental.

## 3.2 MEDIAL POSITION

## 3.2.1 CASES IN WHICH THE ACCENT CAN BE RECONSTRUCTED UNEQUIVOCALLY

* <i>kíta</i>	Ts <i>b-a-ito</i> , At <i>kíta</i> , Tg Cb <i>kíta?</i> 'see' <sup>37</sup>
* <i>útaq</i>	Ts <i>tra-vto</i> , RkBd <i>pua-úta</i> , Ilk <i>úta</i> , Mal <i>muntah</i> 'vomit'
* <i>qútiN</i>	Kn <i>utini</i> , Puy Hutil, CB <i>útin</i> 'penis' <sup>38</sup>

## 3.2.2 CASES WHICH REQUIRE DISCUSSION

In the following cases Formosan languages all reflect a long-vowel root, but Philippine languages reflect a short-vowel root. The Madurese cognates also reflect a short-vowel root, which makes it likely that the long-vowel penult developed in Formosan languages. In any case the PF had a long-vowel penult for these forms.

* <i>bátu</i>	Kn <i>váatu</i> , Ts <i>fatu</i> , At <i>btu-nux</i> , <sup>39</sup> Tg <i>bató</i> , Cb <i>batú</i> 'stone'
* <i>pítu</i>	Kn <i>u-pítu</i> , Ts <i>pitu</i> , Rk <i>pítu</i> , At <i>pitu?</i> , Tg <i>pitó</i> , Cb <i>pitú</i> 'seven'

The following example shows a disyllabic root with a long-vowel penult. The prefix is \**qali-* found in many languages with names of small animals, insects, sea creatures and the like.

* <i>qaNimátek/qaNimátaq</i> <sup>40</sup>	Kn <i>?animətək-a</i> 'creek leech', <i>?alimətək-a</i> 'paddy leech', <i>nimóca?ə</i> 'paddy leech'; Rk <i>Imótkə</i> , Tg <i>limátik</i> , Cb <i>limátuk</i> , Mal <i>halintah</i> 'leech'
--	--

In the following case it is not possible to reconstruct the accent:

* <i>tutu</i>	Kn <i>maká-tutu</i> 'hit with fist'; Ts <i>m-uutu</i> 'strike'; Jav <i>tutu</i> 'pound rice'; Sam <i>tutu</i> 'beat bark for cloth'
---------------	---

## 3.3 FINAL POSITION

## 3.3.1 CASES IN WHICH THE ACCENT CAN BE RECONSTRUCTED UNEQUIVOCALLY

* <i>búut</i>	Kn <i>vuútu</i> , Rk <i>bútu</i> 'squirrel'; Tg <i>bú?ot</i> 'rabbit'
* <i>páqet</i> (PSF)	Ts <i>pó?tə</i> , Bun <i>pa-pahut</i> 'horsefly'

<sup>37</sup>The final glottal stop in Philippine languages is unexplained. Tagalog shows two alternants, one with and one without glottal stop. With certain affixes the root has a glottal stop and with certain others it does not. Tsuchida reconstructs this root with a final glottal stop.

<sup>38</sup>This form is not quoted by Tsuchida.

<sup>39</sup>Atayal *btunux* shows accent shift to the second syllable of the root with the addition of the suffix *-nux*.

<sup>40</sup>Tsuchida reconstructs \**Limatek* and \**LimeCaq*. The Malay *halintah* retains an initial syllable /ha-/ which reflects the first syllable of the prefix \**qaNi-*. Tsuchida's reconstruction with C is based on Kakanavu which has three alternative reflexes of this word. One of the alternant forms in Kakanavu has /c/ instead of /t/. The /c/ in that form is purely a question internal to Kakanavu.



### 3.3.2 CASES WHICH REQUIRE DISCUSSION

In the following case Cebuano has a short-vowel penult (for in the Philippine languages all roots with a reflex of \*e in the penult have a short vowel).<sup>41</sup> The Tsouic forms point to a long vowel in the penult, but Atayal and Rukai show a short vowel. The initial /s/ in some languages is from a prefix, as the meaning of the Atayal form indicates,<sup>42</sup> and the difference in accent between Tsouic and Rukai and Atayal probably has to do with the fact that the first syllable is contracted with the prefix.

\*épat Kn *u-súpatə*, Ts *səptə*, Rk *patə*, Tg *ápat*, Cb *upát* 'four'; At *spat* 'eight'

### 3.3.3 COUNTERCASES

The following form is a counterexample. However, if the Malay or Javanese forms reflect a medial cluster at an earlier stage, this form may still have retained the cluster in PF and still pattern as a long-vowel penult.

\*qe(n)tút Kn *ʔutútu*, Tg *utót*, Cb *utút*, Mal *kentut*, Jav *entot* 'flatulence'

## 4. TRISYLLABIC ROOTS

### 4.1 LONG-VOWEL ROOTS

Trisyllabic roots with a long-vowel penult, reflect \*C.

\*C in initial position:

\*Calínga Kn *caínga* 'ear ornament'; Rk *cngíra*, Cb (dialectal) *talínga* 'ear'

\*Cumqi (PRT) Kn *cimíʔi*, Rk *cmí* 'cheeks'; Ts *cmiʔi* 'temples'

\*CuqélaN Kn *cuʔúanə*, Ts *cərhə*, Cb *tulʔan* 'bone'

\*C in medial position:

\*aCápi (PRT) Ts *copi*, Rk *acapi* 'goby (fish)'

\*amíCi Kn *n-amíci*, Ts *mici*, Pai *s-amci Solanum nigrum*

\*beCékaj (PSF) Kn *vəə-vəcəkai*, Pak *vəcəkad-an* 'middle, centre'

\*beCúgu (PRT) Kn *ma-vəcərəkə* 'satiated'

\*biCúka Kn *civúka*, Ts *cfuʔo* 'stomach'; Rk *bcúka*, Tg *bitúka* 'intestines'

\*jaRáCu (PSF) Kn *carácu*, Ts *trocu*, Rk *gácu* 'body louse'

\*qa(n)Cípa Ts *acipa*,<sup>43</sup> At *qsipa*, Kpp *antípa* 'kind of turtle'

\*qaNíCu Ts *hícu* 'evil spirit'; Rk *alícu* 'kind of holy tree'; Tg *anító* 'idolatry'; Cb *anítu* 'benign spirit'; Mal *hantu* 'evil spirit'

\*qiCéluR Kn *ʔicúuru*, Ts *f-curu*, Tg Cb *itlog* 'egg'

<sup>41</sup>Tagalog *ápat* 'four' developed from a form with the first syllable reduplicated: \**ee-pát*.

<sup>42</sup>The prefix \*s- had a meaning something like 'one'. Literally, \*s-*épat* would mean 'one *épat* – one fold' or the like. The Atayal form came to mean 'two *épat*' or the like by some process we do not understand, but the fact that the Atayal cognate has this meaning makes it likely that \*s- was indeed a prefix. This explanation of the initial /s/ in Formosan languages is preferable to hypothesising another phoneme for PAN (Tsuchida's \*x).

<sup>43</sup>The initial /a/ in Tsou is unexplained. Tsuchida suggests that this word in Tsou is not an inherited form, as is very likely in view of the semantic character of this word as a term for fauna.

## 4.1.1 COUNTERCASES

The following example has a short-vowel penult but shows reflexes of *\*t* rather than *\*C*. It is the kind of word likely to be re-formed by sound symbolism and thus proves nothing.

*\*qatímula* (PSF)      Kn *?atímua*, Ts *timro*, Puy *Hatimura?* 'flea'

## 4.2 SHORT-VOWEL ROOTS

Trisyllabic roots with a short-vowel penult reflect *\*t*.

*\*qa(ta)tabang* (PSF)      Kn *?atavángə*, RkBd Rk *atábngə* 'cockroach'

*\*tinə?un/tənə?un*      Ts *m-ənvənə*, Rk *u-inúnu*, At *tminun*, Mal *tenun* 'weave'; At *tninun* 'fate'<sup>44</sup>

In the following example Tsou reflects a long-vowel root but Kanakanavu reflects a short-vowel root. However, there is a tendency in Kanakanavu for the accent to shift to the penult in short-vowel trisyllabic roots.

*\*taNiu(d,z)* (PSF)      Kn *tanúcu*, Ts *tahzucu*, RkBd *taľuDu* 'mulberry'

In the following case Cebuano reflects a long-vowel penult, but Ilokano and Tagalog have a short-vowel penult:

*\*bituqen/\*bituqan*<sup>45</sup>      Pai *vituqan*, Ilk *bituén*, Cb *bitú?un*, Tg *bituwín* 'star'

The following forms, although they are related, show such a large number of phonological irregularities that they do not allow for the establishment of a form in the proto-language. Clearly they have undergone analogical re-formations.

*\*tagəRang/tagərang*      Kn *takərang* 'sternum'; Ts *t?omga*, Puy *tahərang* 'chest'; Pai *tjagərang* 'trachea, larynx'; Kpp *tagyan* 'ribs'

## 5. MONOSYLLABIC ROOTS

Monosyllabic roots show *\*C*. The occurrence of *\*C* could be explained by a hypothesis that monosyllabic roots consisted of two morae with stress on the second mora. However, the evidence to prove this hypothesis has not yet turned up.

*\*Caú*      Kn *cáau*, Ts *cou*, Rk *i-cóo*, Tg *táqo*, Cb *táwu*<sup>46</sup> 'person'

## 6. CONCLUSIONS

There is a fair number of forms with *\*C* and *\*t* for which the PF accent can be determined unequivocally. The vast majority of these forms confirms the hypothesis that the distribution of *\*C* and *\*t* is determined by the stress pattern of the root. There is only a handful of counterexamples, and many of these turn out to be forms which are not of the sort to provide good evidence for the phonology of reconstructed forms. Phonological developments motivated by accentual phenomena

<sup>44</sup>This form is the direct passive of the verbal root, which tends to reflect the accentual pattern of the root in PF.

<sup>45</sup>This form is not quoted by Tsuchida.

are in any case certain to have exceptions because of the very nature of accent (that is, that stress shift is correlated with morphological features and there is wide scope for analogical shift of stress).<sup>47</sup> There is an equal number of forms for which the evidence for the accentual pattern of the root in PF is not entirely without question. In some of these cases the reconstructed accent is close to certain; in others it is much less certain. None of these forms offer evidence against the hypothesis, but rather they tend to confirm it with great certainty in some cases and with less certainty in others. In my opinion more work on the reconstruction of the accent in PF will add confirmation to this hypothesis.<sup>48</sup> Whatever the case may be, even if we have not proven beyond the shadow of a doubt that the conditions here cited give the environments for the distribution of the two allophones of PAN \*t, \*t and \*C, we must indeed accept that they were allophones and not two contrasting phonemes, for to do otherwise – to accept the existence of a PAN \*C - \*t contrast – would be to accept the far-fetched consequence thereof, namely, that the Philippine languages are more closely related to the Oceanic languages than to the neighbouring languages of Formosa, with which they have so much in common.

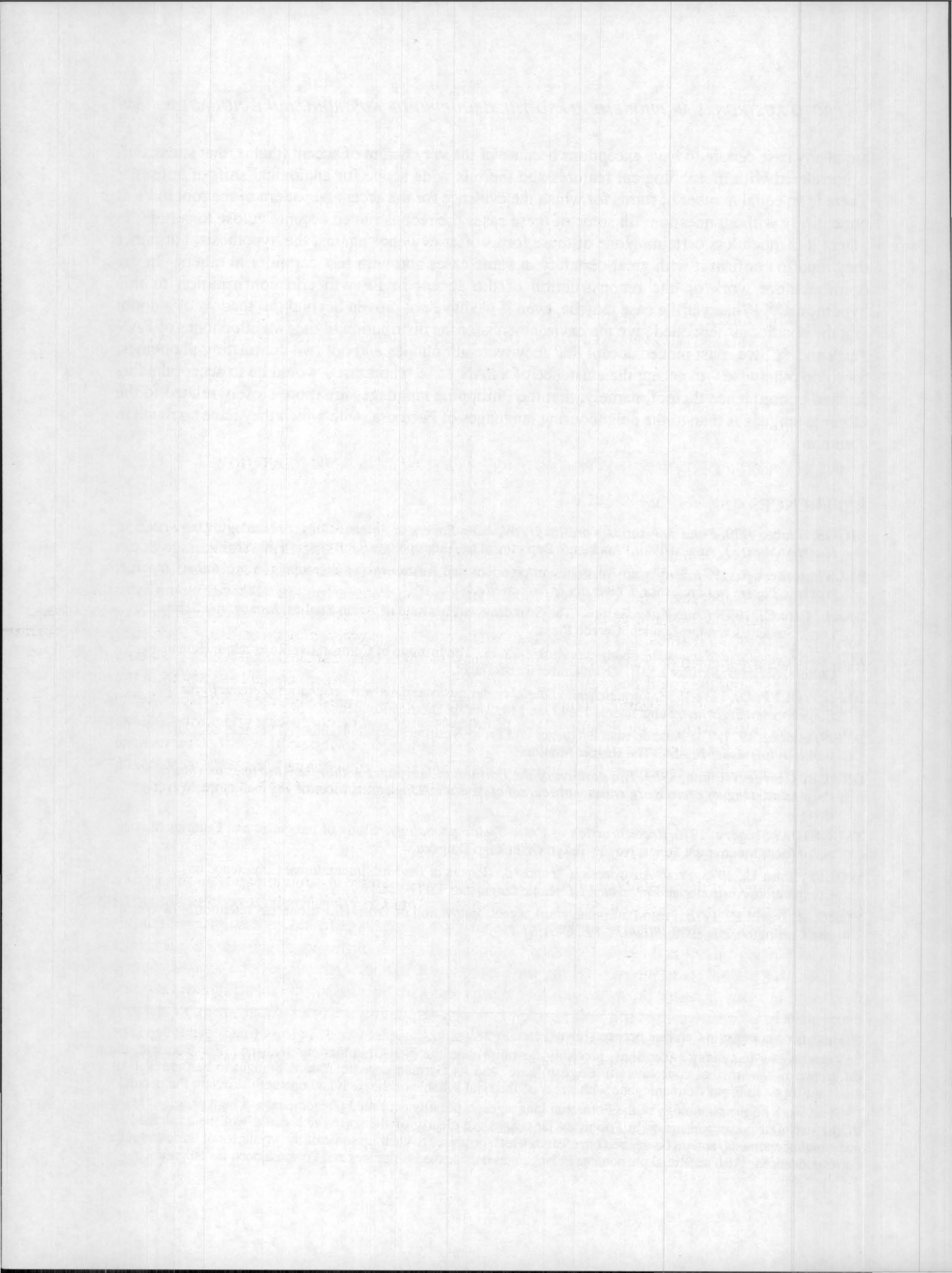
## REFERENCES

- BIGGS, Bruce, 1979, *Proto-Polynesian word list II. Working Papers in Anthropology, Archeology, Linguistics, Maori Studies*, 53, August 1979. Auckland: Department of Anthropology, University of Auckland.
- BLUST, Robert A., 1977, The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report. *Working Papers in Linguistics, University of Hawaii* 9/2:1-15.
- DAHL, Otto C., 1973. *Proto-Austronesian*. The Scandinavian Institute of Asian Studies, Monograph Series 25. Lund: Studentlitteratur; London: Curzon Press.
- 1981, *Early phonetic and phonemic changes in Austronesian*. The Institute of Comparative Research in Human Culture, Series B: Skrifter LXIII. Oslo: Universitetsforlaget.
- DEMPWOLFF, O., 1934-1938, *Vergleichende Lautlehre des austronesischen Wortschatzes. Zeitschrift für Eingeborenen Sprachen*, Supplements 15, 17, and 19. Berlin: Dietrich Reimer.
- DYEN, Isidore, 1971, The Austronesian languages and Proto-Austronesian. In Thomas A. Sebeok, ed. *Current trends in linguistics* 8:5-54. The Hague: Mouton.
- GRACE, George William, 1959, *The position of the Polynesian languages within the Austronesian (Malayo-Polynesian) language family. International Journal of American Linguistics*, Memoir 16. Baltimore: Waverly Press.
- TSUCHIDA, Shigeru, 1976, *Reconstruction of Proto-Tsuoic phonology*. Study of Language and Cultures of Asia and Africa, Monograph Series, No. 5. Tokyo Gaikokugo Daigaku.
- WOLFF, John U., 1975, Proto-Austronesian \*r and \*d. (Papers of the First International Conference on Comparative Austronesian Linguistics.) *Oceanic Linguistics* 13:77-121.
- ZORC, R. David P., 1978, Proto-Philippine word accent: innovation or Proto-Hesperonesian retention? In Wurm and Carrington, eds 1978 [SICAL]. PL, C-61:67-119.

<sup>46</sup>There is no contrast in Cebuano between [áwu#] and [awú#].

<sup>47</sup>Verner's law has many exceptions, probably far more than the Neo-Grammarians realised. For example, the difference in the medial consonants in English 'hare' and its German cognate *Hase* is thought to be the result of alternant forms in Proto Germanic, one with stress on the final syllable of the stem and one with stress on the penult.

<sup>48</sup>Good work on the phonology of the Formosan languages, especially accentual phenomena, is a high priority. Many of the irregular correspondences in Formosan languages are clearly of the sort which could well be explained by accentual phenomena, and this is an area of research which promises to clear up some of the mysteries of these irregular correspondences. Also accentual phenomena in languages outside the Philippines and Formosa need investigation.



## LANGUAGE DECAY AND REVIVALISM: THE ÄÏIWO LANGUAGE OF THE REEF ISLANDS, SANTA CRUZ ARCHIPELAGO, SOLOMON ISLANDS

S.A. WURM

### 1. INTRODUCTION: THE ÄÏIWO LANGUAGE, ITS POSITION AND FEATURES

#### 1.1 THE REEFS-SANTA CRUZ FAMILY

The ÄÏiwo language is a member of the small Reefs-Santa Cruz family of languages which is situated on the Reef Islands and the main island of Santa Cruz in the Santa Cruz Archipelago at the eastern extremity of the Solomon Islands chain. The family, which consists of two sub-families, has three members, one of them the ÄÏiwo sub-family-level isolate, spoken by over 4000 people on most of the Reef Islands and by a large number of Reef islanders who are relatively recent immigrants to northern parts of the main island of Santa Cruz. The other sub-family, called the Santa Cruz sub-family, is wholly located on Santa Cruz Island, and has two members: Santa Cruzan which has about 3500 speakers, and the small Nangu language which has about 250 speakers and is located in the south-western part of the island. Santa Cruzan consists of two sub-languages, Northern Santa Cruzan, also called Natigu ('our language') and Lödäi ('beach') in the north-west and north of Santa Cruz, with about 2300 speakers, and South-western Santa Cruzan in the south-western part of the island with about 1200 speakers. Both sub-languages of Santa Cruzan have several dialects which differ quite markedly from each other, whereas the dialectal differentiation of ÄÏiwo is quite minor.

#### 1.2 NATURE OF THE LANGUAGES OF THE REEFS-SANTA CRUZ FAMILY

The languages of the Reefs-Santa Cruz family deviate profoundly in structure and typology from the six aberrant Austronesian languages spoken in the Santa Cruz Archipelago on the islands of Vanikoro and Utupua, and also from the Polynesian Pilheni language found on the northern islands of the archipelago (Duff and the northern Reef Islands). They also differ equally profoundly from the Austronesian languages of the eastern Solomon Islands and the northern islands of Vanuatu. While they contain many possible Austronesian loan words and a number of peripheral structural features, especially in possessive marking and pronouns, which are undoubtedly of Austronesian origin, it seems obvious that these are due to borrowing. The probable loan words belong to different types of Austronesian, and difficulties are encountered in trying to establish regular sound correspondences



(Peter Lincoln pers. comm.). At the same time, the core grammars of the Reefs-Santa Cruz languages display Papuan features, some of which correspond systematically with comparable features in Papuan languages of the Solomon Islands chain, and further west. It seems justified to regard the languages as originally Papuan languages which have been subject to very strong Austroneisan influence. Prehistoric evidence (Wurm 1978) appears to offer some explanation of how this could have come about.

### 1.3 FEATURES OF THE LANGUAGES OF THE REEFS-SANTA CRUZ FAMILY

The languages, which are of very considerable complexity, have some quite unusual structural and semantic characteristics (Wurm 1987). They have polysynthetic features, and many of their lexical items are composed of small meaningful elements some of which do not occur in isolation but carry specific meanings. A large proportion of their vocabulary is descriptive in nature, and many are verbal nouns consisting of a verb stem with or without affixes, and provided with one of the quite numerous class or gender prefixes found in the languages, especially in Äÿiwo. The languages show evidence of a high degree of preoccupation of their speakers with classifying the world around them into large ranges of categories and divisions. In each of them, there are three separate cross-cutting class systems in the noun phrase, that is, a semantically based noun class system in which the classes are marked by prefixes (Wurm 1981a), a possessive class system and a construct possessive noun class system (Wurm 1981b). The three systems show very little relationship to each other, except that some of the classes in the possessive systems reappear in the construct class systems in the Santa Cruz sub-family languages.

### 1.4 THE NOUN CLASS SYSTEM IN ÄÿIWO

Turning to Äÿiwo, in which the noun class system is more elaborate than in the other languages of the family, it can be observed that the majority of the classes in the semantically based noun class system indicate the nature, and in one case, the location of the items referred to by the nouns carrying the class prefixes. The class prefixes are either found with what appear to be original noun bases or, more frequently, verbs which function as verbal nouns when the class markers are prefixed to them. Many of the bases involved can take different class prefixes of this kind. These classes, of which 39 have been established to date, can be called variable noun classes.

The remaining noun classes in Äÿiwo, of which eight have been found so far, denote the shape, appearance, and the specific nature of the relationship to other things, of the items referred to by the nouns appearing with class prefixes of this kind. The bases to which they are added are very largely original noun bases. Allowing for a few quite spectacular exceptions, most of the noun bases concerned appear only with one particular class prefix of this kind, but sometimes, variable noun class prefixes are found prefixed to the class prefixes of this second kind, which in turn are placed before the noun (or rarely, verb) bases. Perhaps not quite appropriately, the classes denoted by this second kind of class prefix can be called fixed noun classes.



### 1.5 CONCORDANCE IN THE NOUN PHRASE IN ÄËIWO

Sixteen of the variable noun classes are accompanied by concordance phenomena of varying types and extent within the noun phrase. There are indications that until recently, the concordance system in the noun phrase in ÄËiwo was more extensive and complex, and also involved the third person singular pronouns.

### 1.6 CLASSIFICATION OF VERBS IN ÄËIWO THROUGH PREFIXES

ÄËiwo also has a system of classifying verbs through the addition of prefixes to them which indicate modes of action, that is, that an action is carried out using one's hand, or a tool, or a cutting tool, or in the form of a single violent action, etc. (Wurm 1987).

## 2. GRADUAL DECAY OF THE ÄËIWO LANGUAGE

### 2.1 GENERAL REMARKS

The abovementioned structural features have been chosen for discussion because they are the ones most strongly affected in the speech of the younger generation, which has had an increasing exposure to the outside modern world during the last two to three decades. Many of the morphological and semantic characteristics and complexities of ÄËiwo reflect and are linked with features of the traditional culture and world view of its speakers, and the decay and disappearance of the latter may constitute a major reason for the simplification and decay of grammatical complexities in the language. The same factors, as well as increasing pressure from Solomon Islands Pijin and English, are responsible for major changes, and the impoverishment of the vocabulary of the language. Many complex items consisting of individual small elements reflecting traditional indigenous thinking and world view, but no longer properly understood by the younger generation of speakers, are replaced by loans from Solomon Islands Pijin and English. Similarly, large sections of the vocabulary denoting elements, concepts and activities of the traditional culture have increasingly fallen into disuse.

The decay and simplification of the structural features of ÄËiwo will now be illustrated with some examples. However, it will be necessary to say a few words first about the practical alphabet now in use for the language, and employed in providing the examples.

### 2.2 THE ÄËIWO PRACTICAL ALPHABET

For over a century, missionaries of the Church of Melanesia, and a few other persons concerned with languages of the Solomon Islands and adjacent parts of Island Melanesia, have been trying to unravel the structures of the languages of the Reefs-Santa Cruz family, and to devise usable practical alphabets for them. These attempts failed, and whatever was written in the languages and occasionally even printed, is very largely incomprehensible to speakers of the languages today, even to those closely familiar with the present-day practical alphabets used for them.

About a decade ago, the first efforts were made by educated ÄËiwo speakers to reduce their language to writing with the help of outside linguists. This led to the creation of a fully phonemic alphabet devised by Martin MoËiya who was studying for the priesthood, and by Patrick Bwakolo,

who was training to become a secondary school teacher, and who received some linguistic training as well. This alphabet seemed adequate to well-educated and sophisticated Äÿiwo speakers. However, when Patrick Bwakolo began conducting classes which were intended to teach Äÿiwo speakers to read and write their language, considerable difficulties were encountered. These difficulties were due to the presence of a number of subphonemic features in the language, such as whispered and reduced or silent vowels resulting in syllabicity of consonants, and several types of labial or palatal glides separating most vowel sequences, among others. While these features are predictable, the conditioning factors are very complex and the pupils were confused. It was therefore decided to indicate subphonemic features in the alphabet. After several years of practical use in teaching and other situations, a final form of the alphabet was attained. It was found that the present, in part subphonemic, shape of the alphabet was much easier and more useful for the average beginning student of literacy in the language than its phonemic form (Wurm 1986).

In a nutshell, the alphabet is as follows:

VOWELS	CONSONANTS
<i>i</i> <i>u</i>	<i>p</i> <i>t</i> <i>k</i>
<i>e</i>	<i>b</i> <i>d</i> <i>dy</i> <i>g</i>
<i>o</i>	<i>kw</i>
<i>ä</i>	<i>bw</i> <i>gw</i>
<i>a</i> <i>ä</i>	<i>n</i> <i>ny</i> <i>ng</i>
length: doubling <i>aa</i>	<i>ngw</i>
drop and rise of intensity	<i>v</i> <i>s</i>
in long vowel <i>a·a</i>	<i>l</i>
whispered <i>î, û, ê</i>	<i>w</i> <i>y</i>

GLIDES: labial *w*, palatal *y*, whispered palatal *ÿ*

### 2.3 CONCORDANCE IN THE NOUN PHRASE

It was mentioned that 16 of the variable noun classes show various types of concordance in the noun phrase. This concordance occurs with qualitative adjuncts, numerals and possessives (with the latter usually only if possession is emphasised). The concordances shown reflect the use of the language by an average speaker aged over 60 today.

(1) Full concordance within the noun phrase, with qualitative adjuncts, numerals and possessives:

- vä-*    chicken class, e.g. *opwa* 'to be white', *lakî* 'to be small', *eve* 'three', *nugo* 'my (food)'  
 > *väwopwa välakî väÿeve vänûgo* 'my three small white chickens'
- de-*    thing class, e.g. *dekîlingä* 'food', *päko* 'to be good', *polegî* 'six', *na* 'his (food)' >  
*dekîlingä depäko depolegî dena* 'his six (types of) good food'

- u*<sub>1</sub>- flowing liquid class, e.g. *pe* 'to move along', *ki-* continuing or customary action prefix > *ukîpe ulakî uýeve* 'three small streams'
- u*<sub>2</sub><sup>-1</sup> banana class, e.g. *va* 'to be immature', *nyigî* 'one' > *uva unyigî una* 'his one tiny banana'
- bo*<sub>2</sub>- shark-type class, e.g. *bova bolakî bonyigî bona* 'his one small shark (for food)'

(2) Concordance with qualitative adjuncts and possessives (if possession is emphasised), but with numerals the concordance is marked by *pe* (the marker of the person class):

- gi*<sup>-2</sup> masculine with humans class, e.g. *nou* 'my (babies, general)' > *giva gîlakî penyigî ginou* 'my one small male baby'
- si-* feminine with humans class, e.g. *sîva sîlakî penyigî sînou* 'my one small female baby'
- me-* collective humanness class (irrespective of sex), e.g. *meva melakî peýeve menou* 'my three small babies'
- pi-* human adolescent class, e.g. *piva pîlakî penyigî pinou* 'my one small adolescent'

(3) Concordance with qualitative adjuncts: numerals generally have no class prefix (only on rare occasions, the class prefix appears with them). Possessives are provided with the adjunct prefix *mi-* if possession is emphasised.

- ka-* mutability class, e.g. *pûlopwe* 'to be grey', *lâpo* 'to be dirty' > *kapûlopwe kalâpo minou* 'my dirty dust'; *kakîlolo* 'grasshopper' (from *lolo* 'to skip') > *kakîlolo eve minou* 'my three grasshoppers'
- bo*<sub>1</sub>- cavities in wood class, e.g. *lilî* 'to be small', *bomobo* 'rafter', *nugû* 'my (gadgets)' > *bolilî bomobo eve minugû* 'my three short rafters'
- bu-* ground class, e.g. *ngango* 'to be strong', *miyou* 'to be heavy' > *bûngango bûmiyou minugû* 'my heavy clay (used for a purpose)'; *bûngango nyigî* 'one (heap of) clay'
- oli-* cut nut class, e.g. *olîwopwa olîpâko nyigî minûgo* 'my one good cut nut (with white flesh)'
- opo-* house class, e.g. *uko* 'to have aura of tabu', *to* 'my (of immobile things)' > *opokuuko opolilî (opo)nyigî mito* 'my one small church'

(4) Concordance with numerals only. Qualitative adjuncts are preceded by the adjunct prefix *mi-*, and possessive markers are also preceded by the adjunct prefix marker *mi-* if possession is emphasised.

- pe-* person class, e.g. *sîngeda* 'female' > *pesîngeda milakî penyigî minou* 'my one small female person'
- nwa-* fruit class, e.g. *nwanyiga* 'pandanus fruit' > *nwanyiga milakî nwanyigî minou* 'my one small pandanus fruit'

<sup>1</sup>*u*<sub>2</sub>- is derived from the noun base \**u* 'banana'; 'banana' is *nou*, with the prefix of the fixed noun class *no* – of the indistinctly set off parts of whole noun class.

<sup>2</sup>In affixes and bases containing *-i-*, *-u-*, this vowel changes to *-î-* or *-û-* in many instances when the affixes are added to words, or affixes are added to bases containing *-i-*, *-u-*.

(5) No concordance. Qualitative adjuncts are preceded by the adjunct prefix *mi-*, numerals have no prefix and possession markers are preceded by the adjunct prefix *mi-* if possession is emphasised.

The remaining 23 variable noun classes belong to this category, and so do all eight of the fixed noun classes. Example:

*be-* basket class, e.g. *nûmwaañû* 'fish (generic: caught)' > *benûmwaañû milakî nyigî minûgo* 'my one small fish basket'

With younger speakers, there is an increasing tendency, in direct proportion with decreasing age, for fluctuating free variation between the types of concordance described and illustrated under (1)-(4) and what has been mentioned under (5). With teenagers, (5) is used prevalently or almost exclusively. For instance, the examples given above can be observed in the following decaying forms:

- (1) *vä-* : *väwopwa välakî eve vänugo; väwopwa välakî eve nugo; väwopwa milakî eve nugo*  
*de-* : *dekîlîngä mipäko polegî dena*  
*u<sub>1</sub>-* : *ukîpe ulakî eve; ukîpe milakî eve*  
*u<sub>2</sub>-* : *uva nyigî una*  
*bo<sub>2</sub>-* : *bova milakî nyigî bona*
- (2) *gi-* : *giva pelakî penyigî ginou; giva milakî nyigî ginou*  
*si-* : *sîva pelakî nyigî sînou; sîva milakî nyigî sînou*  
*me-* : *meva melakî eve menou; meva milakî eve menou; meva milakî eve penou*  
*pi-* : *piva milakî nyigî pinou; piva milakî nyigî penou; piva gîlakî nyigî penou*
- (3) *ka-* : *kapûlopwe miläpo minou, kapûlopwe miläpo nou; kakîlolo eve nou*  
*bo<sub>1</sub>-* : *bolilî mimobo eve (mi)nugû*  
*bu-* : *bûngango mimiyou (mi)nugû*  
*olî-* : *olîwopwa mipäko nyigî (mi)to*  
*opo-* : *opokuuko mililî nyigî (mi)to*
- (4) *pe-* : *pesîngedä milakî nyigî (mi)nou*  
*nwa-* : *nwanyigä milakî nyigî (mi)nou*

#### 2.4 CLASS PREFIXES

With young speakers, dropping of some of the variable and also of fixed noun class prefixes (in the latter case, only if the same base can appear with several of them) has been observed. For instance in present-day 'correct' Äÿiwo, 'coconut' is *nenû*, and 'coconut palm' *nyanenû* with the tree class prefix *nyä-*. Only in noun phrases can the *nyä-* be dropped. However, young speakers use *nenû* for both 'coconut' and 'coconut palm' in all cases. As another example, the noun base *\*-paa* 'sliver of wood or metal' which does not occur in 'correct' Äÿiwo without noun class prefixes,

appears in the following forms and meanings: with the pointed objects class prefix *nyi-* as *nyipaa* = 'a chip' ('pointed sliver'); with the indistinctly set off parts of a whole class prefix *no-* as *nopaa* = 'pieces of bark beginning to peel off a tree trunk' (i.e. 'slivers indistinctly set off and still attached to the thing to which they belong and of which they form a part'); with the moving away into the distance class prefix *nyo-* as *nyopaa* = 'type of arrow' (i.e. 'sliver-type object which moves away into the distance'); and with the foreign, in particular Polynesian, provenance class prefix *te-* as *tepaa* = 'nail' (i.e. 'sliver of metal of foreign origin'). Young speakers tend to omit the prefixes and only use *paa* by itself in all four of the meanings mentioned.

## 2.5 CONSTRUCT POSSESSIVE NOUN CLASSES

There are a number of construct possessive noun classes in Äÿiwo, for example, a base, background or material construct class marked by *(y)ä* placed between the two nouns involved (e.g. *nwopwa yä nyiivä* = 'stone house'); a purpose construct class indicated by *nä* (e.g. *nupo nä sii* = 'fishnet', 'net for the purpose of catching fish'); a derivation construct class indicated by *lä* (e.g. *nwoÿi lä sapolo* = 'papaya juice', i.e. 'water of papaya'), etc. Young people tend to confuse these class markers and give a preference to *(y)ä*.

## 2.6 COMPLEX VERBAL NOUN FORMS

There is a type of verbal noun in Äÿiwo with which possession is expressed through the adding of the usual subject markers found with fully inflected verb forms, with these subject markers followed by a set of special suffixes which also vary for person. These verbal nouns denote a number of ordinary objects such as 'ladder', 'throat', 'belt', 'appearance', 'heart', etc., for example, *nyekiliÿe* = 'ladder, stairs', made up as follows: location noun class prefix (*nye-*) – continuous or imperfective aspect marker (*ki-*) – 3pl subject indicator (*li-*) – 'to step on something' (*-iiÿe*). 'My ladder' adds *-no* (1sg subject) and *-ngä* (special verbal noun suffix 1sg), that is, 'the place on which I continuously step'; *nye-kû-lu-mw-ä* = 'your (sg) heart', that is, location noun class prefix (*nye-*) – continuous or imperfective aspect marker (*ku-*) – to live (*-lu*) – you (sg) (*-mw*) – special verbal noun suffix 2sg (*-ä*), namely, 'the place where you continuously live'. Young people often cannot handle these verbal nouns and their final suffixes properly and tend to replace some of them by loans from Solomon Islands Pijin, for example, *lata* = 'ladder, stairs', or avoid them through using synonyms such as *nabûlela* = 'ladder'.

## 2.7 CLASSIFICATION OF VERBS THROUGH PREFIXES INDICATING MODE OF ACTION

This feature of Äÿiwo was mentioned in section 1.6. A few examples may be given here:

*vä-* indicates that an action is carried out using one's hand, for example, *vä-dû* 'to beat someone with one's hands'; *vä-gäte* 'to slit something, tearing it by hand'; *vä-gî* 'to push something with one's hands (standing behind it)'; *vä-ginyimä* 'to scream or yell in pain, with arms flailing'.

*tä-* using a tool or involving an object in the action of separating things (for example, using a knife, or opening a basket through separating its lid from it, etc.), as in *tä-ÿi* 'to gather up things into a basket (for example, separating them from the ground)'; *tä-de* 'to slice something (as carrots, with a knife)'; *tä-gäte* 'to slit something using an instrument'.



*lä-* indicates that something is cut with a knife or a cutting tool, for example *lä-ngee* 'to cut open a yam, or to open a bottle, using a knife for these purposes'; *lä-päkau* 'to carve with a knife or a carving tool'; *lä-gäte* 'to cut something in half, using a knife'.

As can be seen from the three lexemes *vä-gäte*, *tä-gäte* and *lä-gäte* given above, the same verb base can appear with a variety of mode of action prefixes.

Other examples of verb prefixes of this kind are:

*po-* indicates that energy is exerted and directed at something in carrying out an action, for example, *po-daŷi* 'to turn one's heel up or stretch it into the air forcefully when standing on one's head'; *po-mubû* 'to lie curled up very tightly'; *po-taa* 'to find something'; *po-vilî* 'to wrap something very tightly around something else'.

*to-* indicates that an action is carried out in a single violent action, for example, *to-gûlo* 'to hit someone or something with a single strong blow, or to stab or spear someone or something'; *to-go* 'to bump into something'; *to-bengî* 'to shut something'; *to-pûle* 'to stab with a spear'; *to-waamûka* 'to make a sudden feint with a stick at someone'; *to-yaa* 'to push a stick against something that is hanging'.

Young speakers tend to mix up these verb prefixes, confusing *lä-* and *tä-* for instance and preferring to use only *tä-*. Similarly, they avoid using *to-* with verbs indicating violent actions which are carried out using one's hand and employ *vä-* instead, for example, *vä-gûlo* = 'to beat someone (repeatedly) with one's hand or a stick' is used by them both in that meaning and for 'hit someone with a single strong blow' which should correctly be *to-gûlo*.

## 2.8 SEMANTICALLY RELATED CONCEPTS INDICATED BY PHONOLOGICALLY SIMILAR LEXEMES

What has been mentioned in section 1.7 may be briefly illustrated here: *tepeke* 'skull of a dead person', *temenge* 'skull of a living person'; *ngegiläve* 'to labour in childbirth', *ngegeläve* 'to moan'; *mwange* 'to bite', *mängä* 'to laugh'; *mwäŷeyaa* 'to bite and pull', *mwäŷeyaa* 'to suck'; *nûmatäpî* 'betel nut', *nunûmotäpî* 'a small coconut which looks a little like a betel nut' (the prefix *nu-* is the marker of the dependent items noun class which indicates that items denoted by nouns belonging to this class are dependent on something else for their existence or their characteristics); *taula* 'to float (freely)', *taulaa* 'to float (tied up, like a canoe)'; *mema* 'to be easy', *memaa* 'to be smooth', etc.

Young speakers are often confused over these items and tend to use only one member of each pair for both concepts, for example, *tepeke* for both 'skull of a dead person' and 'skull of a living person', *taula* for both meanings of 'float', etc.

## 2.9 LEXICAL ITEMS COMPOSED OF SMALL MEANINGFUL ELEMENTS

This feature of Äŷiwo and other languages of the Reefs-Santa Cruz family has been briefly mentioned in section 1.3. Some examples are:

*nye-kû-po-lo* 'volcano' (*nye-* location noun class prefix, *ku-* continuous aspect verb prefix, *po-* 'to burn', *-lo* focusing suffix on verbs indicating that something is very strongly affected by an action, that is, lit. 'a place which burns habitually and strongly affects something, that is, through its eruptions').



*de-lû-po-vi-lî* 'cigarette' (*de-* thing and non-person noun class prefix, *lu-* 3pl subject, *po-* mode of action verb prefix indicating that energy is exerted and directed at something in carrying out an action, *vi-* concept of being bent around something, *-li* focusing suffix on verbs indicating that the object is only indirectly or superficially affected by the action expressed by a verb to which the suffix is added, that is, lit. 'something which they wrap around something else without damaging it, exerting energy in the process, that is, wrapping it tightly').

*pu-pe-lî-lâ* 'to pass someone who goes in the same direction, or to overtake someone' (*pu-* 'to go somewhere', *pe-* 'to move along', *li-* focusing suffix, with *pe-lî* meaning 'to move somewhere touching something', *-lâ* direction suffix indicating movement out of something or out of a given situation, that is, lit. 'going somewhere moving somewhere touching something and moving out of this touching situation').

*nubwa-lâgaa* 'spittle' (*nubwa* 'murky water', *lâgaa* 'his deep breath', that is, lit. 'murky water (expelled by) his deep breath').

*nyi-ÿengî* 'wind' (*nyi-* pointed, protruding and standing out items noun class prefix, *ÿengî* 'to cry, to howl', that is, lit. 'something pointed (and stinging) which makes a howling noise').

*lo-go-taa-lâ* 'to pick out everything (of interest or value) from something' (*lo-* concept of obtaining something, *\*-go-* 'to be in close contact with something (it does not occur in isolation), *taa* concept of finding, *-lâ* direction suffix indicating movement out of something or out of a given situation, that is, 'obtain something through getting in close contact with something in finding objects with (the objects) moving out of this situation').

*ÿyâ-ng(e)-ee* 'to copulate' (*ÿyâ-* prefix indicating interaction between similar or complementing partners, with backward and forward motion involved, *ng(e)-* 'to be at the same place', *-ee* upward slanting movement suffix, i.e. 'two complementary partners interact at the same place with backward and forward motion at a slant').

Young speakers do not use some of the lexical items of this kind, because they no longer understand their compositions. Of the examples given above, some young speakers did not correctly understand *pupelîlâ*, and did not understand *logotaalâ* at all.

### 3. REVIVALISM INVOLVING THE ÄÿIWO LANGUAGE

The Äÿiwo language had reached the level of decay as described in section 2 at the time at which a satisfactory practical alphabet for the language became available. This was accompanied by the increasing availability of texts in the language for reading purposes, with most of them based on materials collected from old speakers still possessing a perfect command of its traditional form and vocabulary, and still very much at home, at least in their memories, in the traditional indigenous culture. At the same time, a dictionary project was started which is still in progress and which, in a prepublication form available to many literate Äÿiwo speakers, contains over 4000 items, many of which have been unknown to young speakers of the language in their forms, meanings and cultural connotations. This has greatly contributed to fixing the vocabulary and the structure of Äÿiwo in a permanent form and to stem the tide of further decay and loss of vocabulary. The teaching of literacy and of features of the language itself at some schools in the Äÿiwo-speaking area has further enhanced the knowledge of the language by young speakers. This encouraging development was very much helped by a resurrection of features of the traditional indigenous culture as a deliberate

policy, such as carving, canoe-building, the production of traditional artefacts and ornaments, traditional dancing, ceremonies and songs, all this accompanied by a revival of the associated vocabularies. Young speakers have begun to reverse the situation discussed above, understand much more of the traditional vocabulary, and are again becoming familiar with the composition of complex lexical items. It is pleasing to see that most of this has come about as a result of the interest and hard work of indigenous speakers of the language, with relatively little outside help.

## BIBLIOGRAPHY

- STICKEL, G. and U. PIEPER, eds, 1985, *Synchronic and diachronic linguistics*. Winter Festschrift. Berlin: Mouton de Gruyter.
- WURM, S.A., 1978, Reef's-Santa Cruz: Austronesian, but ...! In Wurm and Carrington, eds 1978:969-1010.
- 1981a, Notes on nominal classification systems in Äiwo, Reef Islands, Solomon Islands. In Andrew Gonzales and David Thomas, eds *Linguistics across Continents, studies in honor of Richard S. Pittman*, 123-142. Manila: SIL and The Linguistic Society of the Philippines.
- 1981b, The possessive class system in Äiwo, Reef Islands, Solomon Islands. *Papers in New Guinea Linguistics No.21*, 181-209. PL, A-61.
- 1985, Language contacts and special lexical developments. In Stickel and Pieper, eds 1985:961-971.
- 1986, Remarks on some language problems in the Santa Cruz Archipelago, Solomon Islands. In B.F. Elson, ed. *Language in global perspective: papers in honor of the 50th anniversary of the Summer Institute of Linguistics 1935-1985*, 507-523. Dallas, Texas: Summer Institute of Linguistics.
- 1987, Semantics and world view in languages of the Santa Cruz Archipelago, Solomon Islands. In R. Steele and T. Treadgold, eds *Language topics: essays in honour of Michael Halliday*, 439-451. Amsterdam/Philadelphia: John Benjamins.
- WURM, S.A. and Lois CARRINGTON, eds, 1978, *Second International Conference of Austronesian Linguistics: proceedings*. PL, C-61.