

THE  
**Oceanic**  
LANGUAGES

JOHN LYNCH, MALCOLM ROSS  
AND TERRY CROWLEY

routledge | LANGUAGE FAMILY | series

# THE OCEANIC LANGUAGES

The Oceanic Languages form a closed subgroup within one of the world's largest language families, Austronesian. There are between 1000 and 1500 Austronesian languages (estimates vary), with so much structural diversity that they are best handled in two volumes, one on the Oceanic and one on the non-Oceanic Austronesian languages. This division is clear and the grammar sketches in this volume provide a cross-section through the structural diversity of the Oceanic languages which is not available elsewhere. Much of the material is drawn from data collected by the authors and has not been previously published.

The volume contains five background chapters: the Oceanic languages, sociolinguistic background, typological overview, Proto Oceanic, and internal subgrouping. In addition, the volume presents forty-three grammar sketches, selected from the five hundred Oceanic languages spread across a region embracing eastern Indonesia, Melanesia, Polynesia, and Micronesia.

**John Lynch** is Emeritus Professor of Pacific Languages and the former Director of the Pacific Languages Unit at the University of the South Pacific in Port Vila, Vanuatu.

**Malcolm Ross** is Emeritus Professor at the Australian National University.

**Terry Crowley** was a Professor at the University of Waikato in New Zealand.

## ROUTLEDGE LANGUAGE FAMILY SERIES

Each volume provides a detailed, reliable account of every member language, or representative languages of a particular family. Each account is a reliable source of data, arranged according to the natural system of classification: phonology, morphology, syntax, lexis, semantics, dialectology and socio-linguistics. Each volume is designed to be the essential source of reference for a particular linguistic community, as well as for linguists working on typology and syntax.

The Austronesian Languages of Asia and Madagascar

*Edited by Nikolaus Himmelmann & Sander Adelaar*

The Khoesan Languages

*Edited by Rainer Vossen*

The Manchu-Tungusic Languages

*Edited by Alexander Vovin*

The Oceanic Languages

*Edited by John Lynch, Malcolm Ross & Terry Crowley*

The Sino-Tibetan Languages

*Edited by Graham Thurgood & Randy LaPolla*

The Indo-Aryan Languages

*Edited by George Cardona & Dhanesh Jain*

The Iranian Languages

*Edited by Gernot Windfuhr*

The Languages of the Caucasus

*Edited by Alice Harris & Ricks Smeets*

The Bantu Languages

*Edited by Derek Nurse & Gerard Philippon*

The Mongolic Languages

*Edited by Juha Janhunen*

**THE**  
**OCEANIC**  
**LANGUAGES**

John Lynch, Malcolm Ross and Terry Crowley

 **Routledge**  
Taylor & Francis Group  
LONDON AND NEW YORK

First published 2002  
by Curzon Press

First published in paperback 2011  
by Routledge  
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada  
by Routledge  
711 Third Avenue, New York, NY 10017

*Routledge is an imprint of the Taylor & Francis Group, an informa business*

© 2002, 2011 John Lynch, Malcolm Ross and Terry Crowley

The right of John Lynch, Malcolm Ross and Terry Crowley to be identified as authors of this work has been asserted by them in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

*Trademark notice:* Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

*British Library Cataloguing in Publication Data*

A catalogue record for this book is available from the British Library

*Library of Congress Cataloging in Publication Data*

A catalog record for this book has been requested

ISBN: 978-0-7007-1128-4 (hbk)

ISBN: 978-0-415-68155-1 (pbk)

ISBN: 978-0-203-82038-4 (ebk)

Printed and bound in the UK by CPI Antony Rowe, Chippenham and Eastbourne

---

# CONTENTS

---

|                                                                    |           |
|--------------------------------------------------------------------|-----------|
| <i>Preface</i>                                                     | ix        |
| <i>Abbreviations</i>                                               | xii       |
| <i>List of illustrations</i>                                       | xvii      |
| <b>Chapter 1 The Oceanic languages</b>                             | <b>1</b>  |
| 1 The Austronesian family                                          | 1         |
| 2 Geography                                                        | 4         |
| 2.1 Regional profile: Micronesia                                   | 6         |
| 2.2 Regional profile: Polynesia and Fiji                           | 6         |
| 2.3 Regional profile: Melanesia                                    | 6         |
| 3 Demography                                                       | 10        |
| 3.1 Populations                                                    | 10        |
| 3.2 Language and territory                                         | 12        |
| 4 Language contact                                                 | 15        |
| 4.1 Oceanic–Papuan contact                                         | 15        |
| 4.1.1 Papuan Tip OV order                                          | 15        |
| 4.1.2 Allegedly ‘mixed’ languages                                  | 16        |
| 4.2 Contact between Oceanic and non-Oceanic Austronesian languages | 16        |
| 4.3 Contact between different Oceanic languages                    | 16        |
| 4.3.1 Polynesian influence on non-Polynesian languages             | 17        |
| 4.3.2 Direct and indirect inheritance in Rotuman                   | 17        |
| 4.4 Contact with intrusive languages                               | 18        |
| 5 Brief history of research                                        | 19        |
| 5.1 Up until World War II                                          | 19        |
| 5.2 More recent descriptive studies                                | 20        |
| 5.3 More recent comparative studies                                | 20        |
| 5.4 Current state of knowledge                                     | 21        |
| 6 Language names                                                   | 21        |
| <b>Chapter 2 Sociolinguistic background</b>                        | <b>23</b> |
| 1 Socio-cultural background                                        | 23        |
| 2 Language varieties                                               | 24        |
| 3 Vernaculars and lingua francas                                   | 25        |
| 4 Language status                                                  | 28        |
| 5 Written forms                                                    | 30        |
| 6 Oceanic languages into the future                                | 31        |

|                                                     |    |
|-----------------------------------------------------|----|
| <b>Chapter 3 Typological overview</b>               | 34 |
| 1 Phonology                                         | 34 |
| 2 Nouns and noun phrases                            | 35 |
| 2.1 Pronouns                                        | 35 |
| 2.2 Nouns                                           | 37 |
| 2.3 Articles and demonstratives                     | 38 |
| 2.4 Numerals and number-marking                     | 39 |
| 2.5 Adjectives and nominal modifiers                | 40 |
| 2.6 Basic noun phrase structure                     | 40 |
| 2.7 Possession                                      | 40 |
| 2.8 Relative clauses                                | 43 |
| 3 Verbs and verb phrases                            | 43 |
| 3.1 Verbal derivation and inflection                | 43 |
| 3.2 Basic verb phrase structure                     | 45 |
| 3.3 Verb serialisation                              | 46 |
| 4 Clause structure                                  | 49 |
| 4.1 Verbless clauses                                | 49 |
| 4.2 Verbal clauses: core arguments                  | 49 |
| 4.3 Verbal clauses: peripheral arguments            | 51 |
| 4.4 Negative clauses                                | 51 |
| 5 Imperative and interrogative sentences            | 52 |
| 5.1 Imperative sentences                            | 52 |
| 5.2 Interrogative sentences                         | 52 |
| 6 Complex sentences                                 | 53 |
| <b>Chapter 4 Proto Oceanic</b>                      | 54 |
| 1 Theoretical and methodological preliminaries      | 54 |
| 2 The position of Proto Oceanic within Austronesian | 56 |
| 2.1 Reconstructive and notational conventions       | 57 |
| 2.2 Proto Oceanic and Proto Malayo-Polynesian       | 57 |
| 3 A Proto Oceanic grammar sketch                    | 63 |
| 3.1 Phonology                                       | 63 |
| 3.1.1 Phonemes                                      | 63 |
| 3.1.2 Orthography                                   | 66 |
| 3.1.3 Phonotactics                                  | 66 |
| 3.1.4 Stress                                        | 67 |
| 3.2 Nouns and noun phrases                          | 67 |
| 3.2.1 Pronouns                                      | 67 |
| 3.2.2 Nouns                                         | 69 |
| 3.2.3 Articles and demonstratives                   | 70 |
| 3.2.4 Numerals and number-marking                   | 72 |
| 3.2.5 Adjectives and nominal modifiers              | 74 |
| 3.2.6 Basic noun phrase structure                   | 75 |
| 3.2.7 Possession                                    | 75 |
| 3.2.8 Relative clauses                              | 80 |
| 3.3 Verbs and verb phrases                          | 80 |
| 3.3.1 Verbal derivation and inflection              | 80 |
| 3.3.2 Basic verb phrase structure                   | 83 |
| 3.3.3 Verb serialisation                            | 86 |

|                             |                                                |            |
|-----------------------------|------------------------------------------------|------------|
| 3.4                         | Clause structure                               | 86         |
| 3.4.1                       | Verbless clauses                               | 86         |
| 3.4.2                       | Verbal clauses: core arguments                 | 86         |
| 3.4.3                       | Verbal clauses: peripheral arguments           | 87         |
| 3.4.4                       | Negative clauses                               | 88         |
| 3.5                         | Imperative and interrogative sentences         | 89         |
| 3.5.1                       | Imperative sentences                           | 89         |
| 3.5.2                       | Interrogative sentences                        | 89         |
| 3.6                         | Complex sentences                              | 89         |
| <b>Chapter 5</b>            | <b>Internal subgrouping</b>                    | <b>92</b>  |
| 1                           | Theoretical background                         | 92         |
| 2                           | Proto Oceanic and primary subgroups of Oceanic | 94         |
| 3                           | Admiralties subgrouping                        | 99         |
| 4                           | Western Oceanic subgrouping                    | 99         |
| 4.1                         | The Meso-Melanesian linkage                    | 101        |
| 4.2                         | The Papuan Tip linkage                         | 102        |
| 4.3                         | The North New Guinea linkage                   | 106        |
| 4.4                         | The diversification of Western Oceanic         | 106        |
| 5                           | Central/Eastern Oceanic subgrouping            | 108        |
| 5.1                         | The Southeast Solomonian family                | 110        |
| 5.2                         | Utupua and Vanikoro                            | 112        |
| 5.3                         | The Southern Oceanic linkage                   | 112        |
| 5.4                         | The Central Pacific linkage                    | 114        |
| 5.5                         | The Micronesian family                         | 117        |
| 5.6                         | The diversification of Central/Eastern Oceanic | 119        |
| <b>The grammar sketches</b> |                                                | <b>121</b> |
| Kele                        | adapted by Malcolm Ross                        | 123        |
| Mussau                      | by Malcolm Ross                                | 148        |
| Sobei                       | by Joyce Sterner and Malcolm Ross              | 167        |
| Tobati                      | by Mark Donohue                                | 186        |
| Kairiru                     | adapted by Malcolm Ross                        | 204        |
| Takia                       | by Malcolm Ross                                | 216        |
| Arop-Lokep                  | by Lucille S. D'Jernes                         | 249        |
| Jabêm                       | adapted by Malcolm Ross                        | 270        |
| Gapapaiwa                   | by Catherine McGuckin                          | 297        |
| Sudest                      | by Mike Anderson and Malcolm Ross              | 322        |
| 'Ala'ala                    | adapted by Malcolm Ross                        | 347        |
| Bali-Vitu                   | by Malcolm Ross                                | 362        |
| Kaulong                     | adapted by Malcolm Ross                        | 387        |
| Siar                        | by Malcolm Ross                                | 410        |
| Taiof                       | by Malcolm Ross                                | 426        |
| Banoni                      | adapted by John Lynch and Malcolm Ross         | 440        |
| Sisiqa                      | by Malcolm Ross                                | 456        |
| Roviana                     | by Simon Corston-Oliver                        | 467        |
| Kokota                      | by Bill Palmer                                 | 498        |
| Gela                        | by Terry Crowley                               | 525        |

|                                                  |                                     |     |
|--------------------------------------------------|-------------------------------------|-----|
| Longgu                                           | by Deborah Hill                     | 538 |
| Arosi                                            | adapted by John Lynch and Rex Horoi | 562 |
| Buma                                             | by Darrell Tryon                    | 573 |
| Mwotlap                                          | by Terry Crowley                    | 587 |
| Sakao                                            | adapted by Terry Crowley            | 599 |
| Tamabo                                           | by Dorothy Jauncey                  | 608 |
| Raga                                             | abstracted by Terry Crowley         | 626 |
| Vinnavis                                         | by Terry Crowley                    | 638 |
| Port Sandwich                                    | adapted by Terry Crowley            | 650 |
| Southeast Ambrym                                 | adapted by Terry Crowley            | 660 |
| Lamen                                            | by Robert Early                     | 671 |
| Ifira-Mele                                       | by Ross Clark                       | 681 |
| Sye                                              | by Terry Crowley                    | 694 |
| Anejom̃                                          | by John Lynch                       | 723 |
| Cèmuhî                                           | abstracted by John Lynch            | 753 |
| Xârâcùù                                          | abstracted by John Lynch            | 765 |
| Iaai                                             | abstracted by John Lynch            | 776 |
| Ulithian                                         | abstracted by John Lynch            | 792 |
| Puluwatese                                       | abstracted by John Lynch            | 804 |
| Rotuman                                          | adapted by Hans Schmidt             | 815 |
| Nadrogā                                          | by Paul Geraghty                    | 833 |
| Niuafōʻou                                        | abstracted by Robert Early          | 848 |
| Marquesan                                        | abstracted by John Lynch            | 865 |
| <i>Listing of Oceanic languages, by subgroup</i> |                                     | 877 |
| <i>References</i>                                |                                     | 891 |
| <i>Index to Chapters 1–5</i>                     |                                     | 915 |

---

# PREFACE

---

Languages classified as ‘Oceanic’ constitute the largest subgroup of the Austronesian family, itself one of the two largest language families in the world in terms of the number of member languages. Speakers of Oceanic languages live in Melanesia, Micronesia and Polynesia. In Melanesia and Micronesia there are also speakers of non-Oceanic Austronesian languages, and in Melanesia speakers of non-Austronesian languages as well.

There are between 450 and 600 languages classified as Oceanic, depending on whose count one accepts. The exact number is unknown, partly because some of these languages are poorly known, and partly also because it is often difficult to decide where the boundary between language and dialect lies. Most of these languages are spoken by populations that in world terms are very small: no Oceanic language has more than half a million speakers, many are spoken by only a few hundred people, and there is a significant number of Oceanic languages with one hundred speakers or fewer.

Some Oceanic languages have been fairly well known to scholars for well over a century, largely as a result of the descriptive work of missionary-linguists in Polynesia and Fiji in the first half of the nineteenth century. Information about the Oceanic languages of Melanesia and Micronesia for the most part emerged much later. By the second half of the nineteenth century, there was enough information on Oceanic languages of Melanesia in the notes of various missionaries that three compendia of Melanesian Oceanic languages were produced. The first of these, in 1861, was *Die melanesischen Sprachen* by the German scholar H.C. von der Gabelentz. This included sketches of ten languages (none of them from as yet unevangelised Papua New Guinea). It was followed by two highly influential works. R.H. Codrington produced his *The Melanesian languages* in 1885, and S.H. Ray produced *A comparative study of the Melanesian island languages* in 1926. Codrington’s compendium included about three dozen sketches of Melanesian languages for which he was able to gain access to information, while Ray’s volume included about two dozen sketches. The content of these volumes was by and large complementary, which meant that information on a significant number of previously undescribed Oceanic languages was made public.

No scholar of comparative Oceanic languages can fail to make reference to these two works: they are still the only published sources for some of the languages in this subgroup. However, these volumes both have their obvious limitations. None of the sketches could be better than the sources of information upon which they were based. Most of the sketches were written on the basis of Codrington’s and Ray’s own analysis of biblical translations or notes produced by missionaries, and thus it was inevitable that their sketches would contain many errors of both phonetic representation and linguistic analysis.

Although both volumes claim to represent ‘Melanesian’ languages, they were not geographically or genetically representative of the Oceanic subgroup as a whole, or even of those Oceanic languages that are spoken in Melanesia. There is a great concentration of data from the languages of Vanuatu (formerly the New Hebrides) and Solomon

Islands, with the languages of New Caledonia, Papua New Guinea and Irian Jaya being either completely or almost completely ignored. And, of course, they say nothing about Micronesian languages.

We cannot blame Codrington and Ray for this, as these languages were little known at the time. However, the situation today is very different. Large numbers of scholars – both academic linguists and linguistically trained missionaries – have provided vast amounts of new information in the last thirty or forty years on Oceanic languages from all areas. There are now detailed descriptions available of many previously unknown Oceanic languages spoken in Papua New Guinea, Solomon Islands, Vanuatu, New Caledonia, Micronesia and Polynesia. Our understanding of the historical development of these languages has changed immensely as a result.

Of course, there are still many gaps, which means that there is still plenty of scope for fieldworkers planning to produce grammatical descriptions of Oceanic languages. However, it has become obvious that there is quite a lot of information on languages ‘out there’ in places that it is often difficult for practising Oceanic linguists to get access to. Some valuable descriptive material is out of print or in sources that are difficult for libraries to acquire copies of. Some published material is written in a style or in accordance with a model that makes it opaque even to many linguists. There is material that has been distributed to people in particular networks, which others cannot readily gain access to. Given that the majority of practising Oceanists publish in English, material in languages other than English could today also be considered as difficult of access. Finally, there is unpublished (and sometimes only semi-analysed) material in various people’s fieldnotes, including our own. The results of some of our shorter forays into data-gathering would ordinarily not be considered publishable because they would be too short. At the same time, though, we feel that it would be a shame to deny other scholars of Oceanic languages access to these materials.

This volume, then, is an attempt to provide a late twentieth century equivalent to Codrington (1885) and Ray (1926). It is our aim to present an overview of the Oceanic subgroup, and also to provide sufficient phonological and grammatical data to give typologists and comparativists a good idea of the nature of these languages, and of how much typological variety there is in this single subgroup. The references will allow those interested in particular topics, geographical areas or specific languages to delve further.

The book is divided into two major parts. In the five chapters of the first part, we place the Oceanic languages in their geographic, demographic and social context. We deal both with the place of the Oceanic subgroup within the wider Austronesian family, and with the internal subgrouping of Oceanic itself. We provide a typological overview of Oceanic languages, and outline the reconstructed phonology and morphosyntax of Proto Oceanic.

The second part of the book consists of sketch grammars (between 2500 and 10,000 words) of over forty Oceanic languages. This figure represents ten percent or less of the membership of the Oceanic subgroup, and the languages that are represented in these sketches have been chosen according to two main considerations:

- (1) they should represent major genetic or geographical groupings within the Oceanic subgroup as a whole; and
- (2) they should be languages for which information is relatively difficult for practising Oceanists to get hold of.

It should be noted that some of the better known Oceanic languages are not sketched in this volume – languages such as Tolai, Motu, Paamese, Lenakel, Tongan, Samoan,

Māori, Standard Fijian, Hawaiian, Ponapean, Mokilese, etc. Descriptions of these are readily available (see the appendix); and in addition these languages are widely referred to in the general discussions in Chapters 1–5. We have instead deliberately included here sketch grammars of languages which are not well known at all in the linguistic literature. Thus instead of Motu as a representative of the Central Papuan subgroup, for example, we have a sketch of little-known ‘Ala‘ala; instead of an eastern Polynesian language like Tahitian or Hawaiian or Māori, we have a sketch of Marquesan; and so on. Our aim in doing this is to make data on more Oceanic languages accessible. The interested reader can use the bibliography and the appendix to discover where descriptions of better known languages can be found.

Each sketch follows an identical template, which is also adhered to in the chapters dealing with the typological overview of Oceanic languages (Ch. 3) and reconstructed Proto Oceanic (Ch. 4). The length of particular sketches was determined by a variety of considerations. If a language is typologically of particular interest, or if it represents a geographical or genetic grouping that has until now been relatively poorly described, and if we had access to reliable data, then we have aimed to provide sketches of about 10,000 words. Other languages, however, have been described in sketches of between 2500 and 5000 words. The languages are presented in a roughly northwest to southeast order (which corresponds reasonably accurately to the general direction of Oceanic settlement), with Admiralties languages first, then those belonging to the Western Oceanic linkage, then Central/Eastern Oceanic languages. (A full listing of all Oceanic languages, by genetic affiliation and location, is included in the appendix.)

For many of these languages, the sketch in this volume is the first grammatical treatment to appear in print. In other cases, previous research is made more widely available. As a result, there are three categories of authorship represented in the various sketches. If the author's name appears in the usual way beneath the title, this means that the writer of that sketch did original research on that language and that the sketch is not based on anybody else's work. A sketch that is ‘adapted by’ one of the compilers is based on previously published or unpublished grammars; however, either the ‘adapter’ has reinterpreted (parts of) the original source, or has supplemented the information in the grammar with additional research, or both. Finally, sketches that are ‘abstracted by’ one of us basically represent the writer's summarisation of somebody else's work, presented according to our standard template but with little or no fresh input. When we have based a sketch on somebody else's work, we have indicated our sources within the body of the sketch itself.

In the preparation of this book, we have benefited from various kinds of assistance from many people. As well as the authors or co-authors of various sketches, we would like to thank the following for their assistance with particular sketches: Bruce Waters, Nikolaus Himmelmann and Mark Donohue (Takia), Karen Rowe (Siar), John Brownie (Mussau) and Ulrike Mosel (Taiof). In addition, the following either provided comments on the first part of the volume or assisted in other ways: Ross Clark, Robert Early, Jeff Marck, Peter Murgatroyd, Meredith Osmond, Andrew Pawley, Matthew Spriggs, and Holger Warnk.

JOHN LYNCH  
*Port Vila*  
*Vanuatu*

MALCOLM ROSS  
*Canberra*  
*Australia*

TERRY CROWLEY  
*Hamilton*  
*New Zealand*

---

# ABBREVIATIONS

---

## MORPHEME-GLOSS ABBREVIATIONS

Morpheme boundaries are indicated by a hyphen (-), the boundary between a clitic and another morpheme by an equals sign (=). Infixes are enclosed in <...>.

: separates distinct components of portmanteau morphemes; e.g. 2SG:NEG:FUT.

. is used when a single morphological category or lexical item requires two or more words to express its function or meaning; e.g. IMM.FUT (immediate future), NON.FEM (non-feminine), last.night, take.care.of, etc..

The convention for person-number-marking is 1SG, 2DL, but 1INC:DL.

|           |                                 |
|-----------|---------------------------------|
| 1         | first person                    |
| 1EXC      | first person exclusive          |
| 1INC      | first person inclusive          |
| 2         | second person                   |
| 3         | third person                    |
| A         | transitive subject (i.e. agent) |
| ABIL      | abilitative                     |
| ABS       | absolutive                      |
| ACC       | accompaniment                   |
| ACV       | accusative                      |
| ADV       | adverbial                       |
| AF        | affection                       |
| AG        | agent(ive)                      |
| AL        | alienable                       |
| ALL       | allative                        |
| ANAPH     | anaphoric                       |
| ANIM      | animate                         |
| ANTIDESID | antidesiderative                |
| AOR       | aorist                          |
| APPOS     | appositive                      |
| ART       | article                         |
| ASSOC     | associative                     |
| B         | boundary marker                 |
| BEN       | benefactive, beneficiary        |
| BITE      | action involving the teeth      |
| BR        | basic root                      |
| C         | complementiser                  |

---

|           |                           |
|-----------|---------------------------|
| CARD      | cardinal (numeral prefix) |
| CAUS      | cause, causative          |
| CF        | contrafactual             |
| CHAR      | characteristic            |
| CHNG.ST   | change-of-state           |
| CL        | classifier                |
| CM        | class-marking suffix      |
| COL       | collective                |
| COM       | comitative                |
| COMP      | comparative               |
| COMPL     | completive                |
| COND      | conditional               |
| CONST     | construct suffix          |
| CONT      | continuous, continuative  |
| COP       | copula                    |
| CTRST     | contrastive               |
| CUSTOM    | customary possession      |
| CV        | connecting vowel          |
| D         | dependent                 |
| DAT       | dative                    |
| DEF       | definite                  |
| DEIC      | deictic                   |
| DELIM     | delimiter                 |
| DEM       | demonstrative             |
| DER       | derivational prefix       |
| DESID     | desiderative              |
| DET       | determiner                |
| DIM       | diminutive                |
| DIR       | direction(al)             |
| DISJ      | disjunctive               |
| DIST      | distant, distal           |
| DIST.FUT  | distant future            |
| DIST.PAST | distant past              |
| DISTR     | distributive              |
| DL        | dual                      |
| DQT       | direct quote marker       |
| DRINK     | drinkable possession      |
| DUB       | dubitative                |
| DUR       | durative                  |
| EMOT      | emotive, emotional        |
| EMPH      | emphatic                  |
| ERG       | ergative                  |
| ES        | echo-subject              |
| EV        | excrecent vowel           |
| EXC       | exclusive                 |
| EXHORT    | exhortative               |
| FAC       | facilitative              |
| FAM       | familiar                  |
| FEM       | feminine                  |

|           |                                                           |
|-----------|-----------------------------------------------------------|
| FOC       | focal, focus                                              |
| FOOD      | food possession                                           |
| FREQ      | frequentative                                             |
| FUT       | future                                                    |
| GEN       | general possession                                        |
| GNV       | genitive                                                  |
| GOAL      | goal                                                      |
| HABIT     | habitual                                                  |
| HAVE      | ownership                                                 |
| HORT      | hortative                                                 |
| HYP       | hypothetical                                              |
| ID        | identical reference                                       |
| IMM       | immediate                                                 |
| IMM.FUT   | immediate future                                          |
| IMP       | imperative                                                |
| IMPF      | imperfective                                              |
| INAN      | inanimate                                                 |
| INC       | inclusive                                                 |
| INCEP     | inceptive                                                 |
| INCH      | inchoative                                                |
| INCORP    | incorporative                                             |
| INDEF     | indefinite                                                |
| INDEF.FUT | indefinite future                                         |
| INST      | instrument(al)                                            |
| INT       | intentional                                               |
| INTENS    | intensifier                                               |
| INTER     | intermediate                                              |
| IRR       | irrealis                                                  |
| IT'S      | presentative                                              |
| ITER      | iterative                                                 |
| ITR       | intransitive                                              |
| JUICE     | possession of suckable items                              |
| L         | loose                                                     |
| LIG       | ligature                                                  |
| LINK      | linker                                                    |
| LOC       | locative                                                  |
| LOC.PRO   | locative proform                                          |
| MASC      | masculine                                                 |
| MNR       | manner                                                    |
| MR        | modified root                                             |
| MULT      | multiple subject/object                                   |
| NEC       | necessitive                                               |
| NEG       | negative; NEG1/NEG2/NEG3 used for discontinuous negatives |
| NEG.PERF  | negative perfective ('not yet')                           |
| NEUT      | neuter                                                    |
| NMV       | nominative                                                |
| NOM       | nominaliser                                               |
| NON.FEM   | non-feminine                                              |
| NON.FUT   | non-future                                                |

---

|          |                                |
|----------|--------------------------------|
| NON.PRES | non-present                    |
| NON.SP   | non-specific                   |
| NON.SUBJ | non-subject                    |
| NPM      | nominal phrase marker          |
| NSG      | non-singular                   |
| NSP      | non-specific                   |
| NUM      | numeral particle               |
| O        | transitive object              |
| OBJ      | object(-marker)                |
| OBJ.CL   | object classifier              |
| OBL      | oblique                        |
| ORD      | ordinal                        |
| ORN      | ornamental                     |
| P        | proper                         |
| PALL     | palliative                     |
| PART     | partitive                      |
| PASS     | passive                        |
| PAST     | past                           |
| PC       | paucal                         |
| PERF     | perfect(ive)                   |
| PERS     | personal                       |
| PL       | plural                         |
| PLANT    | plant possession               |
| POL      | politeness marker              |
| POSR     | possessor                      |
| POSS     | possessive (unspecified)       |
| POSS.A   | A-possession                   |
| POSS.O   | O-possession                   |
| POSTP    | postposition                   |
| POT      | potential                      |
| PREC     | precedentive                   |
| PRED     | predicator, predicate-marker   |
| PREP     | preposition                    |
| PREPV    | prepositional verb             |
| PRES     | present                        |
| PRO      | clausal proform                |
| PROC     | processive (action in process) |
| PROHIB   | prohibitive                    |
| PROSP    | prospective                    |
| PROX     | proximate, proximal            |
| PURP     | purpose, purposive             |
| Q        | interrogative                  |
| QUANT    | quantifier                     |
| R, REAL  | realis                         |
| REAS     | reason                         |
| REC      | reciprocal                     |
| REC.PAST | recent past                    |
| REDUP    | reduplication                  |
| REF      | referential                    |

|        |                           |
|--------|---------------------------|
| REFL   | reflexive                 |
| REL    | relative clause marker    |
| REP    | repetitive                |
| RESULT | resultative               |
| S      | intransitive subject      |
| SBJN   | subjunctive               |
| SEMI   | semi-alienable            |
| SEQ    | sequential                |
| SG     | singular                  |
| SIM    | similative                |
| SP     | specific                  |
| STAT   | stative                   |
| STEP   | action involving the feet |
| SUB    | subordinator              |
| SUBJ   | subject(-marker)          |
| SV     | stem vowel                |
| SW     | switch reference          |
| TAM    | tense-aspect-mood         |
| TEMP   | temporal marker           |
| TERM   | terminative               |
| TL     | trial                     |
| TPC    | topic(aliser)             |
| TR     | transitive                |
| UNSP   | unspecific                |
| VAL    | valency-changer           |

#### **GEOGRAPHIC, SUBGROUP AND PROTOLANGUAGE ABBREVIATIONS**

|       |                                         |
|-------|-----------------------------------------|
| CEOc  | Central/Eastern Oceanic                 |
| PCEMP | Proto Central/Eastern Malayo-Polynesian |
| PCP   | Proto Central Papuan                    |
| PEMP  | Proto Eastern Malayo-Polynesian         |
| PEOc  | Proto Eastern Oceanic                   |
| PMP   | Proto Malayo-Polynesian                 |
| PNG   | Papua New Guinea                        |
| POc   | Proto Oceanic                           |
| WOc   | Western Oceanic                         |

#### **OTHER ABBREVIATIONS**

|      |           |
|------|-----------|
| k.o. | kind of   |
| sp.  | species   |
| s.o. | someone   |
| s.t. | something |

---

# LIST OF ILLUSTRATIONS

---

## LIST OF MAPS

|      |                                                                 |     |
|------|-----------------------------------------------------------------|-----|
| 1.1  | The Austronesian family and major Austronesian language groups  | 3   |
| 1.2  | The boundaries of the Oceanic subgroup                          | 5   |
| 1.3  | Micronesia                                                      | 7   |
| 1.4  | Polynesia and Fiji                                              | 8   |
| 1.5  | Island Melanesia                                                | 9   |
| 1.6  | The New Guinea mainland and its offshore islands                | 11  |
| 1.7  | Discontinuous languages – Oro Province, PNG                     | 13  |
| 1.8  | Discontinuous languages – Motu, PNG                             | 13  |
| 5.1  | Possible primary subgroups of Oceanic                           | 95  |
| 5.2  | Admiralties and Western Oceanic subgroups                       | 100 |
| 5.3  | Subgroups within the Meso-Melanesian linkage                    | 103 |
| 5.4  | Subgroups within the Papuan Tip Linkage                         | 105 |
| 5.5  | Subgroups within the North New Guinea linkage                   | 107 |
| 5.6  | Major subgroups within Central/Eastern Oceanic                  | 109 |
| 5.7  | Subgroups within the Southeast Solomonian family                | 111 |
| 5.8  | Subgroups within the Southern Oceanic linkage                   | 113 |
| 5.9  | Some subgroups within the Central Pacific linkage (after Marck) | 115 |
| 5.10 | Subgroups within the Micronesian family                         | 118 |

## LIST OF FIGURES

|     |                                                     |    |
|-----|-----------------------------------------------------|----|
| 1.1 | Higher-order Austronesian subgroups                 | 4  |
| 4.1 | Subgrouping of some of the Central Papuan languages | 55 |
| 4.2 | Some sound changes in some Central Papuan languages | 55 |

---

# THE OCEANIC LANGUAGES

---

The phrase ‘Oceanic languages’ refers to the 450 or so languages which are members of the Oceanic subgroup of the Austronesian language family. In this chapter, we briefly describe that wider family, and then examine the geographical range and demographic characteristics of the languages of the Oceanic subgroup itself. We also pay attention to language contact in Oceania, and briefly outline the history of research into these languages.

## 1 THE AUSTRONESIAN FAMILY

The Austronesian language family is one of the largest in the world in terms of membership, with some 1200 member languages; only the Benue-Congo family in Africa has more. Languages of this family cover an enormous area: they stretch from Malagasy (Madagascar) in the west to Easter Island in the east, and from Taiwan and Hawai‘i in the north to New Zealand in the south (although languages of other families are also spoken in parts of this area).

The existence of wide-ranging linguistic relationships in the Asia-Pacific area was first recognised over 300 years ago. The first semi-formal statement of the relationship between Polynesian languages and Malay and its closer relatives was made by Hadrian Reland in 1708 on the basis of Polynesian wordlists collected by Jacob Le Maire in 1615. The early history of Austronesian studies is recounted in Ray’s landmark study of Melanesian languages (see Preface; Ray 1926:19–25). Ray writes, ‘... in the table at the end of the account of Cook’s second voyage [Cook (1777)] Mr Anderson drew attention to the striking resemblance of the Polynesian numerals to those of the Malay Archipelago and Madagascar.’ It was not long before the Austronesian language family was firmly established, by Lorenzo Hervas y Panduro, who devoted to this subject the last five volumes of his twenty-one volume *Idea dell’ Universo*, published between 1784 and 1787. Ray (1926:19–20) writes of Hervas:

His ‘Catalogo delle lingue’ affirmed the close relationship of the languages west and east, and in the ‘Aritmetica delle Nazioni’ he gives a table showing the agreement of the numerals in Cook’s specimens and those of the Marianas, Philippines, Java, Madagascar and Malay. ... He considered that only two language stocks were represented in the Pacific. One was the mother speech of the black races [of Melanesia] and the other the Malayan. According to Hervas, the latter included the languages of the Malay Peninsula, the Maldives, Madagascar, the Sunda Isles, Moluccas and Philippines, with the languages eastward to Easter Island.

Subsequent studies further elucidated the links between the languages of insular and peninsular Southeast Asia on the one hand, and those of Polynesia on the other. Wilhelm von Humboldt provides a scholarly comparative study of some of them in his

*Über die Kawi-Sprache auf der Insel Java* (1836). The geographically intervening languages of Melanesia and Micronesia, however, posed a problem, as Hervas had remarked: they were spoken by people of a quite different physical type, and they were not initially recognised as being related either to each other or to the languages west or east of them. With hindsight, the reasons for this initial omission are not surprising: (a) within Melanesia there is a large number of non-Austronesian (or Papuan) languages; (b) many of the Austronesian languages of Melanesia and Micronesia have changed rather more radically than languages of other Austronesian-speaking regions, thus obscuring the relationship between them; and (c) in any case, many of the languages of Melanesia and Micronesia were completely, or almost completely, unknown to scholars at this time.

Because of this, the family was initially known, and is sometimes still referred to, as the Malayo-Polynesian family, although that term is now reserved for one of the first-order subgroups of Austronesian. Schmidt (1906:59) attributed the invention of the term 'Malayo-Polynesian' to Wilhelm von Humboldt (1836), and this has often been repeated in the literature. It seems, however, that the term was probably first used in print by the linguist Franz Bopp in 1841 (Ross 1996d). The word 'Austronesian' (*Austro*- 'south'; *-nesia* 'island group') was coined by Schmidt (1899:245) because he felt that 'Malayo-Polynesian' excluded by implication the Austronesian languages of Melanesia and Micronesia.

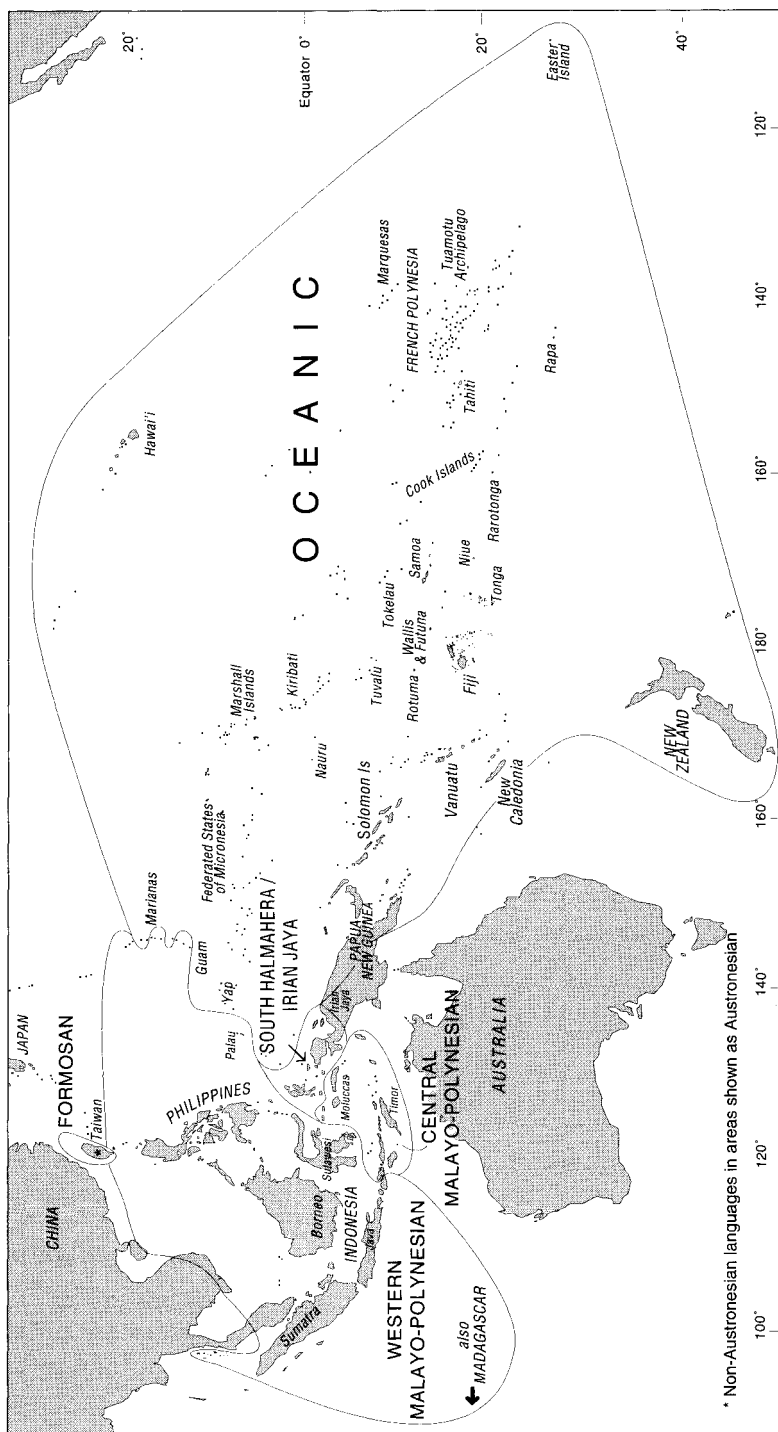
The first serious indications that at least some Melanesian and Micronesian languages belonged to the Austronesian family came in the middle of the nineteenth century. Grace (1976b:57) writes:

... most of the [Melanesian] languages on which information first became available, both in eastern Melanesia and the vicinity of New Guinea, were in fact Austronesian. Latham in his 1847 discussion found no indications of a 'fresh class of languages' (i.e., other than Austronesian) in the Melanesia–New Guinea area.

This is mildly surprising in view of Hervas' observation that such languages did exist. But the chapter of errors continued and Friedrich Müller (1876–88), surveying the world's languages, set up a grouping he called 'Papua-Sprachen' ['Papuan languages'], with two members: Numfoor (Irian Jaya) and Nengone (Loyalty Islands). Both were subsequently shown to be Austronesian: Numfoor by Kern (1885), Nengone by Schmidt (1899).

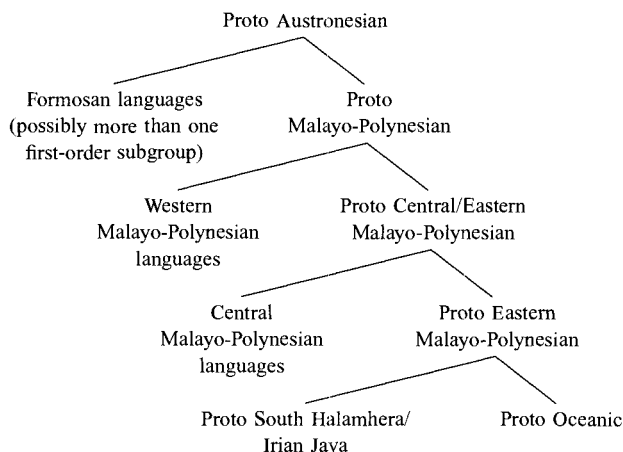
By this time, however, the existence of non-Austronesian languages in what was then British New Guinea had been established in publications by Ray (Ray 1893, 1895). Schmidt (1899), following up on Ray's research, reported similar languages in what was then German New Guinea. Thus, as Grace remarks, the distribution of Austronesian languages in the area was roughly established soon after the turn of the century.

Despite these efforts, the exact nature of the relationships of the languages of Melanesia, Micronesia and Polynesia – both with each other and with the languages of Southeast Asia – remained unclear until the comparative work of Otto Dempwolff, who showed that these languages belonged to a single subgroup of Austronesian, deriving from the intermediate protolanguage *Urmelanesisch* or Proto Melanesian (Dempwolff 1937:190–194). This theory – known now in the literature as the Oceanic Hypothesis – derives from Dempwolff's pioneering studies, and Dempwolff's *Urmelanesisch* is now referred to as Proto Oceanic. (Despite once being referred to as the Oceanic 'hypothesis', there is today no longer any dispute about the validity of the Oceanic subgroup of Austronesian.)



\* Non-Austronesian languages in areas shown as Austronesian

MAP 1.1 THE AUSTRONESIAN FAMILY AND MAJOR AUSTRONESIAN LANGUAGE GROUPS



**FIGURE 1.1** HIGHER-ORDER AUSTRONESIAN SUBGROUPS

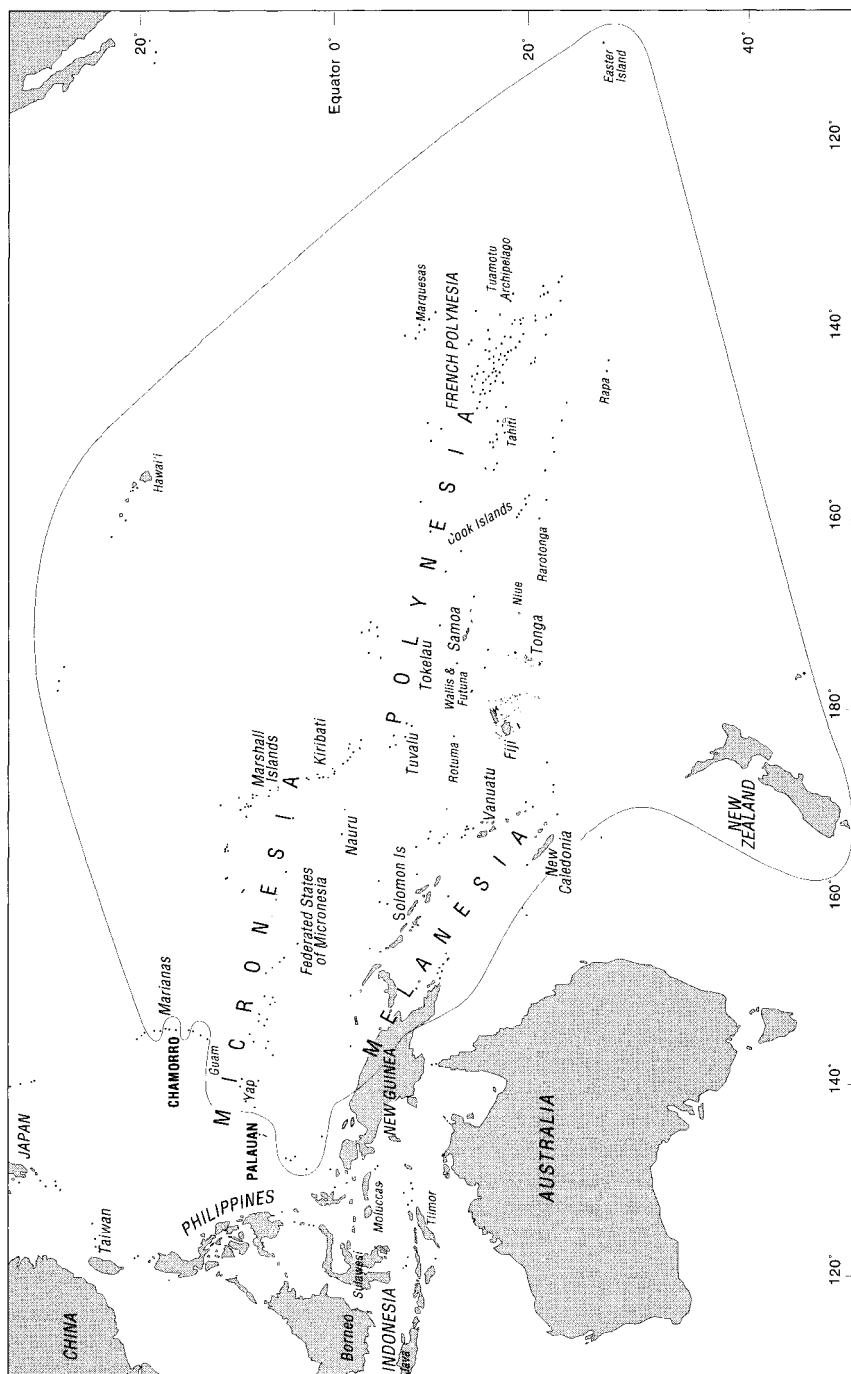
A widely accepted higher-order branching from original Proto Austronesian, following a number of publications by Blust (e.g. Blust 1984), is shown in Fig. 1.1. The geographical distribution of Austronesian and of its major subgroups can be seen in Map 1.1. The innovations which define the Oceanic subgroup are discussed in Chapter 5.

## 2 GEOGRAPHY

The area occupied by speakers of Oceanic languages has been traditionally divided into three sub-regions or 'culture-areas' – Melanesia in the west, Micronesia in the north, and Polynesia in the east – on the basis of a number of geographical, socio-cultural, physical and linguistic factors. This division is somewhat sweeping and superficial (as witness the countless, and generally futile, arguments over whether Fiji belongs with Melanesia or Polynesia) but it is a useful one nevertheless – at least as far as marking Polynesia and Micronesia off from the rest of the region is concerned.

With two exceptions (in Micronesia), all Austronesian languages spoken east of a line drawn roughly along 130° East longitude north of the Equator and along 138° East longitude south of the Equator belong to the Oceanic subgroup. This includes all of Polynesia and Micronesia, and most, but not all, of the Melanesian region, as illustrated in Map 1.2.

However, nothing is as simple as it seems. Two languages spoken in geographic Micronesia – Palauan and Chamorro – are Austronesian but not Oceanic. All Polynesian languages are Oceanic, but the languages of the Polynesian subgroup do not correspond neatly with the geographical area of Polynesia: a number of languages – known as the Polynesian Outliers – which are descended from Proto Polynesian are found geographically in parts of Melanesia and Micronesia. And the Melanesian region, as defined by other (non-linguistic) criteria, contains both Oceanic and non-Oceanic Austronesian languages, as well as seven hundred or so languages belonging to a number of different non-Austronesian families, which are generically (and perhaps somewhat misleadingly) labelled with the cover term 'Papuan'.



MAP 1.2 THE BOUNDARIES OF THE OCEANIC SUBGROUP

### **2.1 Regional profile: Micronesia**

Micronesia is characterised in the extreme east and the extreme west by basically one language per island or per island-group, and in the central area – the Caroline Islands – by quite complex dialect chaining. Schütz (1972:91), in referring to the Fijian dialect-chain, has talked about linguists ‘fall[ing] into the trap of language and dialect counting’, and the Carolines present a classic example of this: Bender and Wang (1985) note that three different linguists have divided this complex continuum into three, seven, and eleven languages respectively!

We have already mentioned the non-Oceanic languages Palauan and Chamorro. Yapese is an Oceanic isolate (Ch. 4, §2). Two Polynesian Outliers are also spoken in Micronesia: Kapingamarangi and Nukuoro (see Mao 1.3). The remaining languages of Micronesia (somewhere between nine and seventeen, according to how one ‘counts’ languages), belong to a subgroup of Oceanic known as Micronesian (Ch 4, §5.5).

### **2.2 Regional profile: Polynesia and Fiji**

Polynesia (see Map 1.4), as the most recently settled large Oceanic-speaking area, corresponds most closely to the characterisation of one language per island or per island-group, often with vast distances between one group or language and the next. Because of distances and time-depths, dialect-chaining is not common in Polynesia, where nineteen languages are now spoken. In Fiji, on the other hand, there is a complex chain of over thirty dialects, with the extreme western and extreme eastern varieties being mutually unintelligible (see Geraghty 1983).

Polynesian is a low-level subgroup of Oceanic, whose closest relatives are the languages of Fiji and Rotuma. All the languages of geographical Polynesia (the ‘Polynesian Trangle’: see Map 1.4) belong to this subgroup, as do fifteen Polynesian Outlier languages spoken in Melanesia or Micronesia (Ch. 5, §5.4). These languages represent back-migrations of Polynesian speakers which occurred some centuries after the initial settlement of Polynesia (which is estimated at having taken place about 1000 BC).

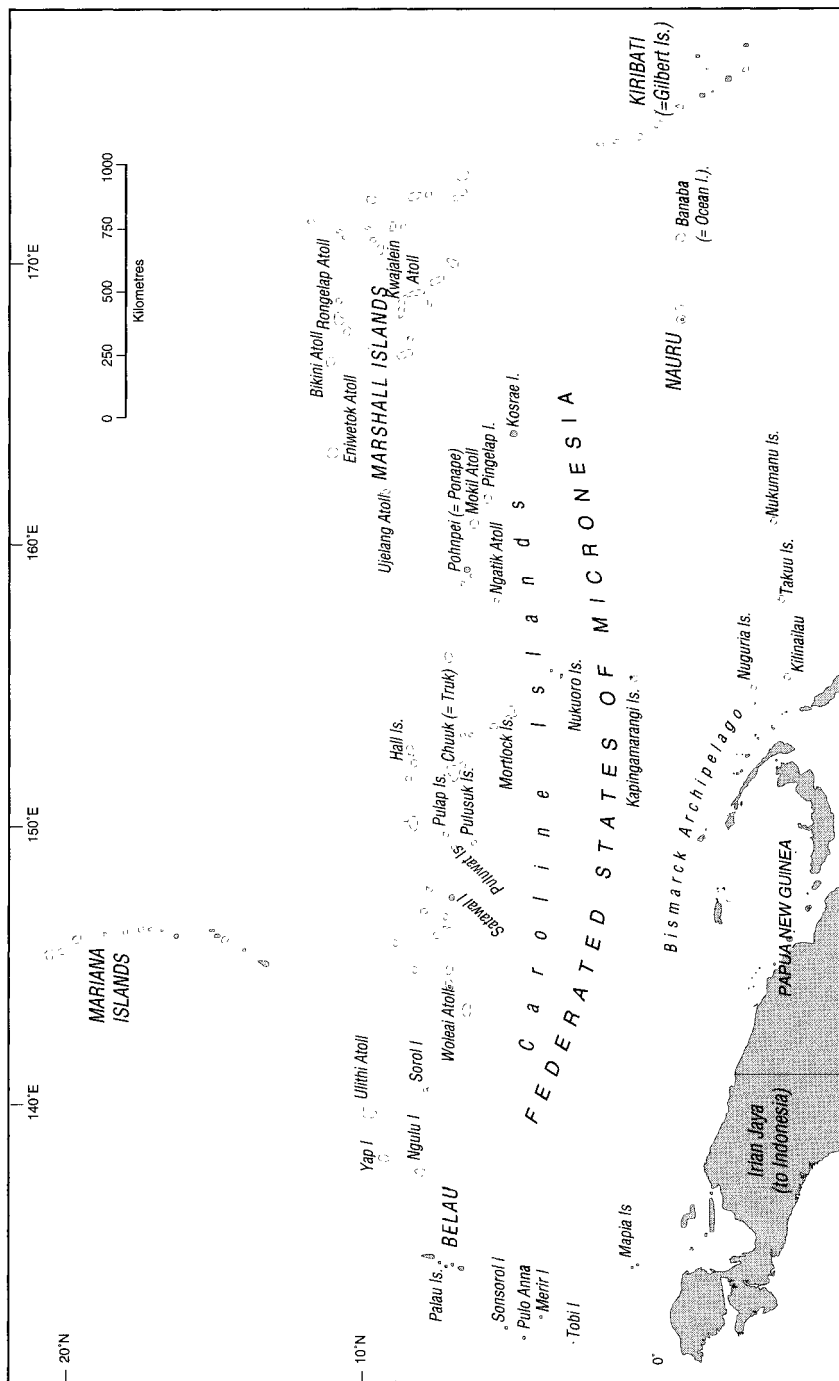
### **2.3 Regional profile: Melanesia**

Melanesia includes what is referred to as Island Melanesia – New Britain, New Ireland and Bougainville (within Papua New Guinea), Solomon Islands, Vanuatu and New Caledonia (Map 1.5). Also included in this area is the mainland of the large island of New Guinea and its numerous small offshore islands.

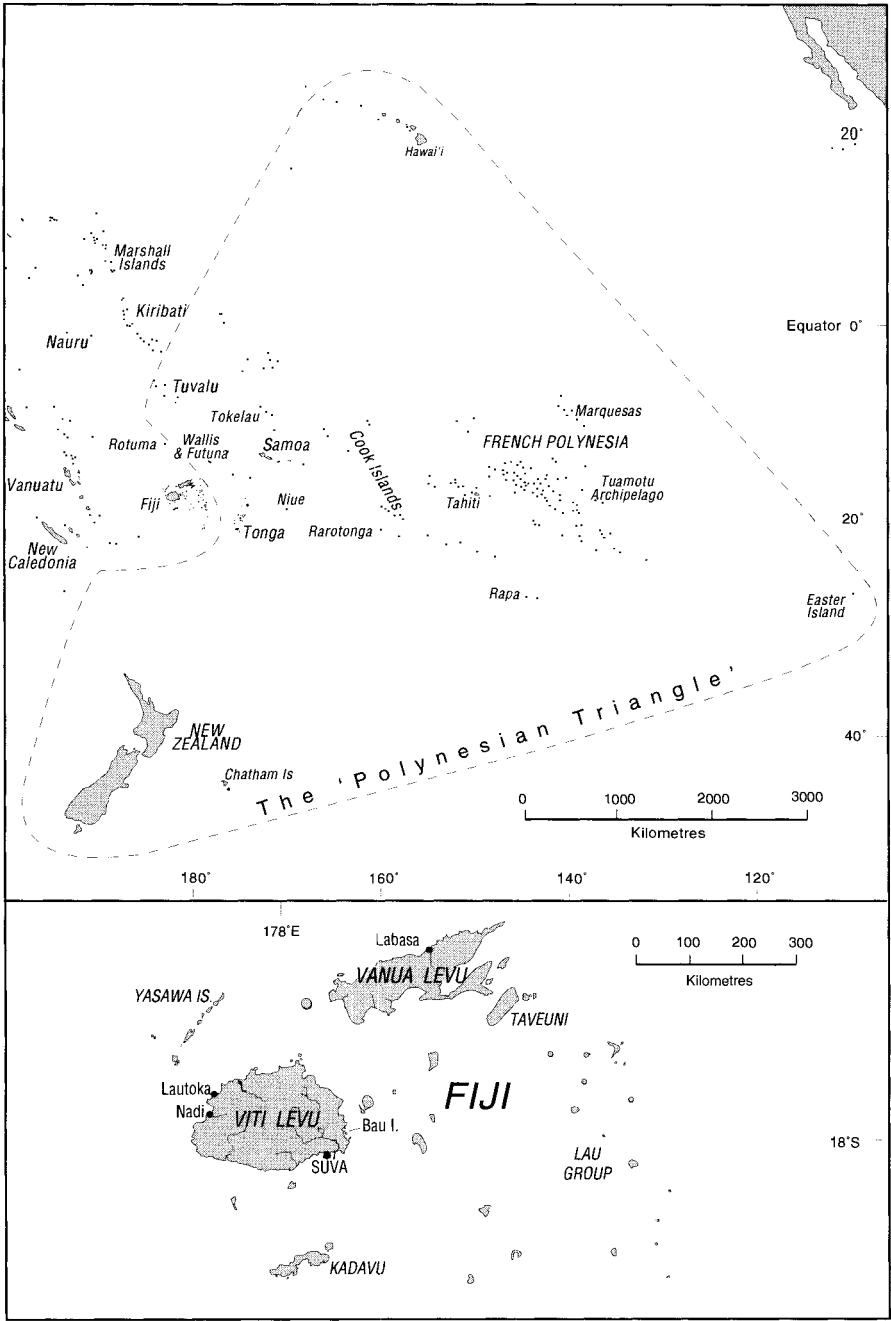
Melanesia is quite different from Micronesia and Polynesia. The situation in Island Melanesia is typified by many languages per largish island. At the same time, there are a number of Papuan languages spoken in the western part of Island Melanesia, often interspersed with Oceanic languages (see Map 1.6).

The linguistic and demographic difference between Melanesia and other parts of Oceania can be illustrated by a couple of comparisons:

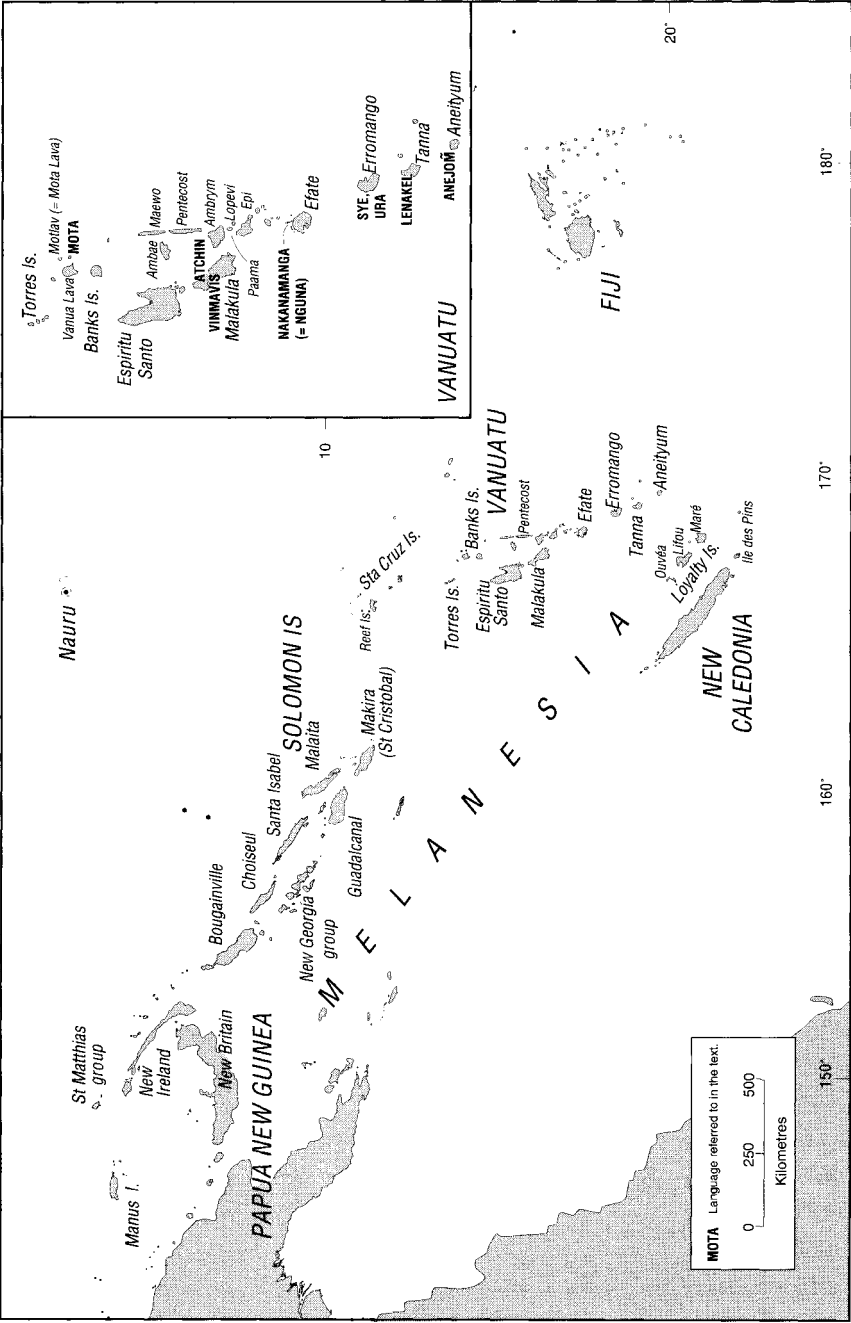
- (1) The islands of Samoa and American Samoa in Polynesia have a land area of just over 3000 km<sup>2</sup>, and the population of about 250,000 speaks just the one language: Samoan. The three islands of Malakula, Ambrym and Paama in central Vanuatu have a slightly smaller total land area than the two Samoas and about one-quarter of the population: nevertheless, thirty-three languages are spoken on these three islands.



MAP 1.3 MICRONESIA



MAP 1.4 POLYNESIA AND FIJI



MAP 1.5 ISLAND MELANESIA

- (2) The land areas of Fiji (with perhaps two or three languages) and Hawai'i (with just one) are each slightly larger than that of Vanuatu. Yet Vanuatu, with a population about half the size as that of the indigenous population of Fiji, has slightly over 100 languages.

The approximate number of Oceanic languages in each of the larger islands (including small offshore islands) or island-groups of Island Melanesia is given in Table 1. The figures in parentheses indicate the number of Polynesian Outliers spoken in that area. Thus, for example, in Bougainville there are fourteen Oceanic languages altogether, of which three are Polynesian Outliers.

About 170 Oceanic languages are spoken on the mainland of New Guinea or on small islands offshore. (We include here the languages of Manus and the other islands in the Admiralty Islands, where about thirty languages are spoken.) These are distributed thinly along the north coast, around the 'tail' of Papua, and for some distance along the south coast. In only three areas do they encroach more than a few kilometres inland: almost 200 km up the Markham Valley west of the city of Lae, and 50 km or so inland in the Rigo and Bereina areas, respectively 60 km east and west of the PNG capital Port Moresby. Despite numbering nearly two hundred, Oceanic languages are heavily outnumbered by the seven hundred or so Papuan languages spoken in the New Guinea area (see Map 1.6).

There is no 'Melanesian' subgroup of Oceanic (Ch. 4, §2). The languages of Melanesia belong to three or more first-order subgroups of Oceanic, one of which also includes the lower-level Micronesian and Polynesian subgroups. (See Map 5.1 in Chapter 5.) However, it is sometimes convenient, especially when discussing typology (Chapter 3), to speak of 'western Melanesia', the area occupied by two of these first-order subgroups and comprising New Guinea, New Britain, New Ireland, Bougainville, and western Solomon Islands (Choiseul, New Georgia, Santa Ysabel).

3 DEMOGRAPHY

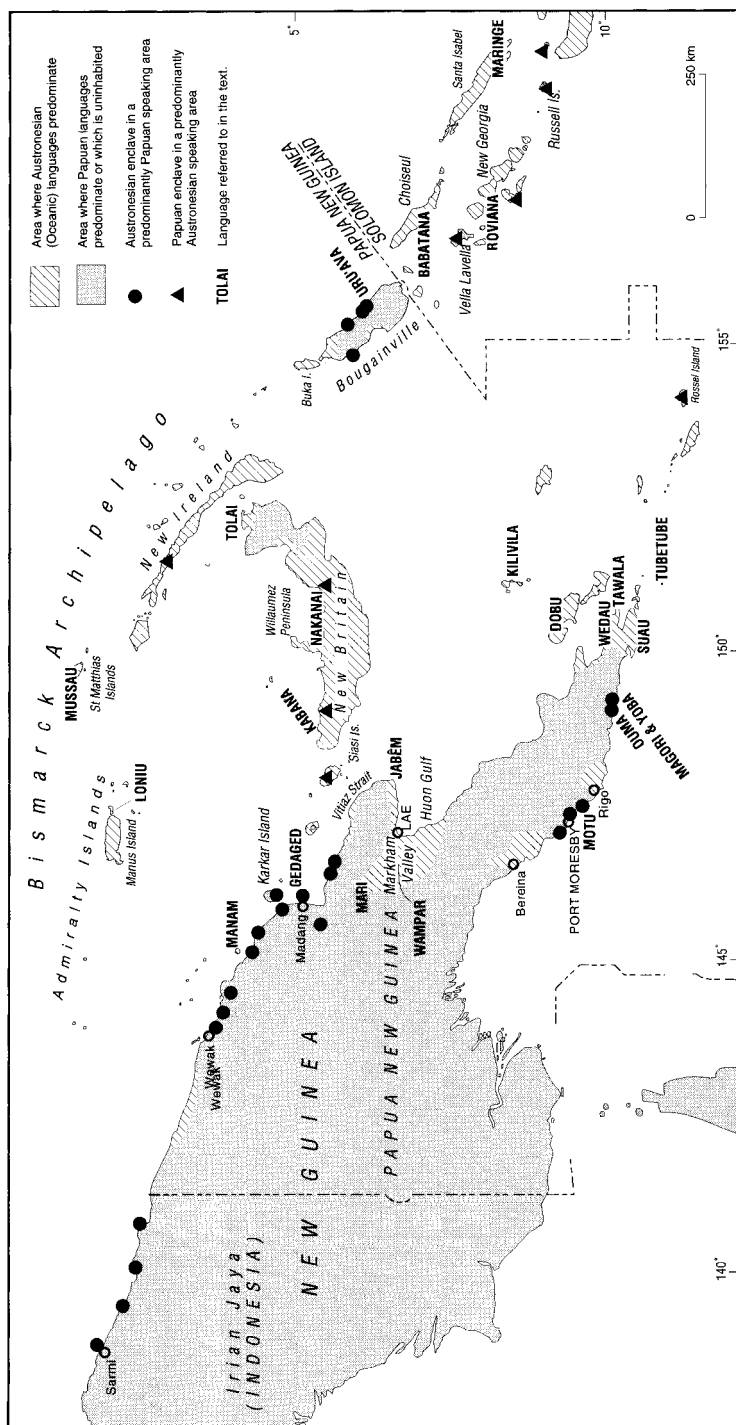
The vast area occupied by less than two million speakers of Oceanic languages necessitates a brief discussion of the linguistic demography of the region. (More detail can be found in Wurm and Hattori's 1981 monumental language atlas of the Pacific region.)

3.1 Populations

Populations speaking individual Oceanic languages vary enormously. Probably the largest Oceanic language, with well over 300,000 speakers, is what is popularly known as 'Fijian'. In reality this is two languages, due to the lack of mutual intelligibility at the

TABLE 1: OCEANIC LANGUAGES IN ISLAND MELANESIA

|                 |     |     |
|-----------------|-----|-----|
| New Britain     | 14  |     |
| New Ireland     | 22  |     |
| Bougainville    | 14  | (3) |
| Solomon Islands | 56  | (6) |
| Vanuatu         | 105 | (3) |
| New Caledonia   | 28  | (1) |



extreme ends of a dialect continuum. Other Oceanic languages with sizeable populations include Samoan, with over 250,000 speakers, and Tolai (New Britain; Map 1.6), Kiribati (Micronesia) and Tongan (Polynesia), each with around 100,000 speakers.

At the other end of the scale are a number of moribund languages (Ch. 2, §6), each with fewer than ten speakers at the last recording (since which time some may have died out). These include Ouma, Yoba, Bina (Papua), Uru'ava (Bougainville), and Ura (Erromango, Vanuatu) (Maps 1.5 and 1.6). Indeed, a small number of Oceanic languages whose existence was noted at the time of early European contact have since become extinct. On the island of Erromango just referred to, for example, at least five languages were spoken 150 years ago; today, apart from the moribund Ura, only the Syle language survives (with almost 1500 speakers).

In this context, averages are probably meaningless, but they can give some indication of the smallness of the populations of Oceanic languages, especially in the western part of the region. The average size of a language (including both Oceanic and Papuan languages together) in PNG and Solomon Islands is about 4000 speakers; this falls to about 2000 in New Caledonia, and to about 1500 in Vanuatu.

While no area of Melanesia can be described as 'typical', we will take the Solomon Islands situation as an illustration of the range of sizes of populations speaking Oceanic languages. Tryon and Hackman's (1983) classification of the languages of Solomon Islands identifies fifty-six Oceanic languages spoken in that country (in which seven Papuan languages are also spoken). Of these fifty-six languages, only two had populations of over 10,000 speakers in 1983 (the North Malaita dialect chain with 13,500 and Kwara'ae, also spoken on Malaita, with 12,500). The populations of Oceanic languages in Solomon Islands (based on 1983 figures) are given in Table 2.

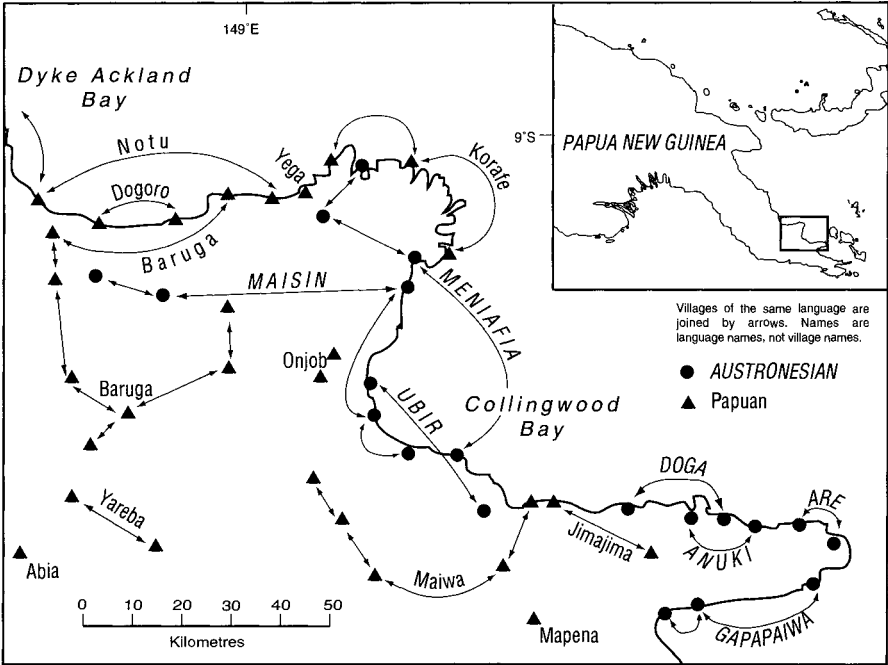
### 3.2 Language and territory

The majority of Oceanic languages occupy their own continuous territory (though of course there are enclaves or dispersed individuals speaking many Oceanic languages in cities, towns and government stations all over the Pacific). In some cases, this 'continuous' territory incorporates vast expanses of open ocean. Kiribati provides an (admittedly extreme) example: 'the total ocean area over which the islands are distributed measures, by various estimates, from 3.5 to 5 million square kilometers', yet the land area is less than 700 km<sup>2</sup> (Bunge and Cooke 1984:277).

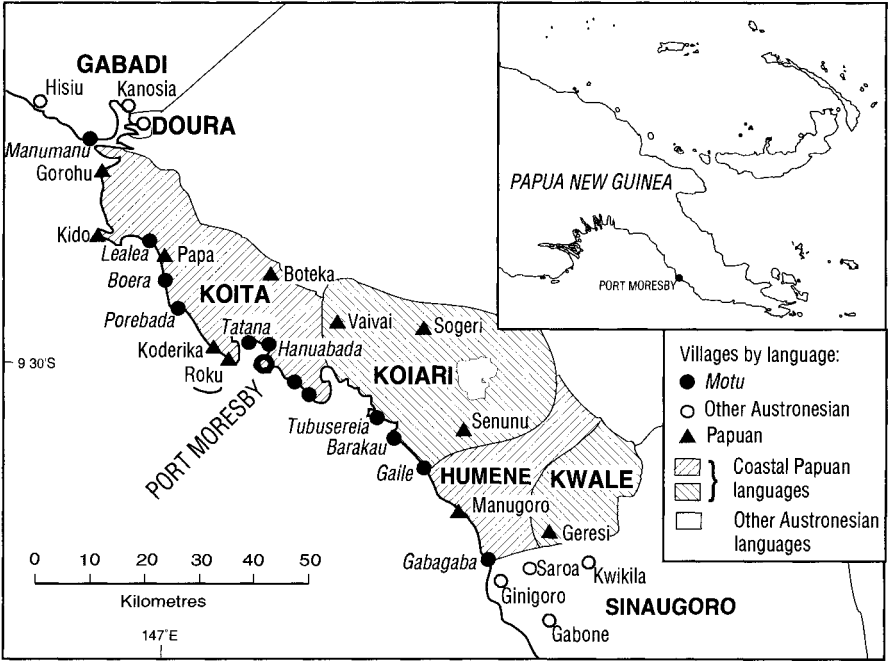
There are, however, two significant types of exceptions to this general statement about 'territorial' languages. The first concerns what we might refer to as 'discontinuous languages'. As a result of fairly recent in- or out-migration, though not over very great

**TABLE 2: POPULATIONS OF OCEANIC LANGUAGES IN SOLOMON ISLANDS**

| Range of population | Number of languages |
|---------------------|---------------------|
| over 10,000         | 2                   |
| 5000-10,000         | 9                   |
| 1000-5000           | 23                  |
| 500-1000            | 8                   |
| 100-500             | 8                   |
| fewer than 100      | 6                   |



MAP 1.7 DISCONTINUOUS LANGUAGES – ORO PROVINCE, PNG



MAP 1.8 DISCONTINUOUS LANGUAGES – MOTU, PNG

distances, the territory of a number of Oceanic languages is interrupted by speakers of one or more other related or unrelated languages. Invasions, resettlement of part of a population, and flight from enemies are three of the causes of these discontinuities. A number of examples could be adduced from different parts of Oceania. Perhaps the most striking case – a result of the aggressive warlike behaviour in the past of the Papuan-speaking Orokaiva people – can be seen in the Collingwood Bay area of PNG (see Map 1.7).

The Motu language of the Port Moresby region of PNG is also a highly discontinuous language. Motu-speaking villages are interspersed with villages occupied by speakers of the Papuan language Koita, as shown in Map 1.8. The largest Motu-speaking village, Hanuabada ('big village'), is actually a fairly recent conglomeration of five smaller villages, one of which is Koita-speaking. The pre-Motu probably lived further inland than they do today, and were pushed down to the coast by the expansion of their Papuan-speaking neighbours. Here, however, they formed a symbiotic relationship with Koita-speakers, involving trade and intermarriage.

The second type of exception to the general rule that Oceanic languages occupy a continuous territory relates to major overseas migrations – recent and not so recent – of significant numbers of speakers of Oceanic languages. The Polynesian Outliers provide a good example of a reasonably ancient 'overseas migration' (see Ch. 5, §5.4 and Map 5.9).

In much more recent times, significant numbers of Samoans and Micronesians have migrated to the United States, while many speakers of Polynesian languages have migrated to New Zealand; in the case of some Polynesian languages, like Niuean, there are more speakers in New Zealand than in the ancestral island; and there are about 20,000 Cook Islanders in the Cook Islands and another 20,000 in New Zealand.

The Kiribati language is perhaps the most extreme example of this type. Kiribati-speakers who once occupied phosphate-rich Ocean Island have been resettled in Fiji. In addition, the severe population pressure on land and on food resources in most of the other atolls of Kiribati has meant that many Kiribati-speakers have had to move to Solomon Islands, Nauru, and other Pacific countries, as well as to the previously unoccupied Christmas Island.

Internal rather than overseas migrations have also taken place in colonial times for a variety of reasons, resulting in discontinuous language areas. Volcanic eruptions on Ambrym and Lopevi in Vanuatu a generation ago saw the permanent resettlement of one village of Southeast Ambrym speakers to a new village on the island of Efate, while the entire population of Lopevi (where a dialect of Paamese was spoken) has been permanently shifted to a new location on Epi. In both of these new locations, the original languages are being maintained.

In many areas where there were originally thinly scattered populations living in inland hamlets, colonial administrators and Christian missionaries encouraged the development of larger coastal villages. In some places, such population movements have drastically altered the original geographical distribution of languages. In southern New Ireland, for example, there was enforced resettlement from hamlets in the mountains to the coast, leaving the original language area unoccupied. Post-colonial developments in southern Malakula described by Charpentier (1982) have seen most of the inland population dispersed into a number of different coastal villages. Some originally contiguous inland languages are now spoken by only very small populations in their original territory, with larger populations speaking the same language in perhaps several different and sometimes geographically widely separated coastal villages, often as linguistic minorities in those villages along with speakers of one or more other languages.

## 4 LANGUAGE CONTACT

Given the demographic characteristics of languages in Oceania, it is not surprising that there has been contact of various kinds between them. In the western part of the region, there are many languages with small populations and small territories. An area of just a few square kilometres may therefore house a number of distinct languages. The area covered by individual languages in the eastern part of the region is often much larger (though, as we have said, much of this ‘territory’ is sea). However, people’s seafaring skills were correspondingly greater in Eastern Oceania, and the sea was more a vehicle for, rather than a barrier to, inter-language contact. (Since colonial times, sea travel has tapered off, and even disappeared altogether in some areas.)

Below we examine contact between Oceanic languages and (a) Papuan languages, (b) non-Oceanic Austronesian languages, (c) other Oceanic languages, and (d) languages intrusive to the region.

### 4.1 Oceanic–Papuan contact

As noted earlier, the western part of the region occupied by Oceanic-speakers is also occupied by seven hundred or so Papuan languages, which themselves belong to a number of different families (for further information, see Wurm ed. 1975, Wurm 1982, and especially Foley 1986).

Oceanic and Papuan languages do not have neat and discrete geographical distributions. In coastal areas, for example, we find Oceanic languages interspersed with Papuan languages (as already discussed), and although Oceanic languages predominate in the islands, there are nevertheless many insular Papuan languages as well.

Given that many Western Oceanic languages are geographically contiguous with Papuan languages, and that most languages of both groups are spoken by small populations, the potential for sociolinguistic contact between Oceanic and Papuan languages is considerable. Lexical copying, for example, occurs frequently, and in both directions. In a number of parts of the region, however, contact has resulted in widespread bilingualism, and in this section we wish to discuss a couple of cases of Oceanic–Papuan contact which have resulted in rather dramatic changes.

#### 4.1.1 *Papuan Tip OV order*

The typical Oceanic (indeed, the reconstructed Proto Austronesian and Proto Oceanic) preferred constituent order is Verb–Object (VO), and oblique phrases are typically introduced by prepositions. Papuan languages, on the other hand, are typically Object–Verb (OV) languages with postpositions. Yet most of the fifty or so Oceanic languages of the Papuan Tip linkage (Ch. 4, §4.2) in the south of the New Guinea mainland have OV order and postpositions. The change from VO to OV order is highly unusual in universal terms. This almost certainly resulted from influence on Proto Papuan Tip (or a couple of its early descendants) by one or more Papuan languages.

Some of the languages of the Madang area on the north coast of New Guinea have not only acquired OV order and postpositions: they have also developed special sentence-medial verb forms matching those in neighbouring Papuan languages.

#### 4.1.2 Allegedly 'mixed' languages

Contact between Oceanic and Papuan languages has also been responsible for a number of cases where different linguists have been in dispute as to the family affiliation of a particular language. Three such cases, discussed in some detail by Lynch (1981), are briefly sketched here:

- (1) Magori (along with three moribund languages), spoken on the south coast of New Guinea, has undergone such heavy lexical influence from the Papuan language Mailu, speakers of which subjugated neighbouring populations in pre-colonial times, that it was classified as Papuan by both Ray (1938) and Capell (1962a). It was only as a result of a more detailed investigation of both morphology and lexicon by Dutton (1976) that it became clear that Magori and its neighbours are in fact Oceanic languages.
- (2) Maisin (see Map 1.7) was classified by Strong (1911) as Oceanic with heavy Papuan grammatical and lexical influence. In the immediately following article in the same journal, Ray (1911) classified it as Papuan, with some Oceanic influence! In basically agreeing with Ray, Capell (1976:571) nevertheless felt that Maisin was 'definitely a case in which a true mixture has taken place'. Lynch (1977b) and Ross (1996a) have shown that the language is in fact Oceanic, but that bilingualism has led to substantial grammatical change, whilst word tabooing (the habit of avoiding words similar to the names of the dead; cf. Ch. 2, §2) has caused a great deal of lexical copying from Korafe, a neighbouring Papuan language.
- (3) Three languages spoken in the extreme southeast of Solomon Islands – Äiwo in the Reef Islands and Santa Cruz and Nanggu in the Santa Cruz group – have been classified as (a) Papuan, with heavy Oceanic admixture (Wurm 1976, 1978), and (b) Oceanic, with heavy Papuan admixture (Codrington 1885, Lincoln 1978). Tryon and Hackman (1983) concur with Wurm on this issue.

#### 4.2 Contact between Oceanic and non-Oceanic Austronesian languages

Yapese, spoken at the western end of Micronesia, sits between non-Oceanic Palauan and the Oceanic languages of the Caroline Islands. Until recently it was the one Austronesian language which had defied classification as either Oceanic or non-Oceanic. It is now fairly clear that Yapese is Oceanic, the descendant of an early immigrant language from Melanesia which has borrowed extensively from nearby Palauan and perhaps another unidentified non-Oceanic source, as well as taking on a heavy admixture from the rather different Micronesian Oceanic languages of the Carolines (Ross 1996b).

#### 4.3 Contact between different Oceanic languages

In all parts of the region, there has also been contact between speakers of different Oceanic languages. In many cases where the languages are very closely related or have very similar phonological and morpho-syntactic histories, such contact is difficult to identify. The more different or distantly related two languages are, however, the easier it is to establish if there has been significant contact between them (as in the case of Yapese Oceanic–Oceanic contact). Some of the more notable cases of such contact are noted below.

#### 4.3.1 Polynesian influence on non-Polynesian languages

The Polynesian Outliers, as we mentioned above, represent relatively recent back-migrations from Polynesia westward into Melanesia and Micronesia. Because the phonological histories of Polynesian languages are quite different from those of the languages to the west, Polynesian contact is often relatively easy to identify.

Migrating Polynesians brought with them cultural or technological complexes or items which did not exist, or which had been lost, in Melanesia and Micronesia. For example, the original settlers of New Caledonia either did not bring the pig with them or, if they did, they ate it to extinction soon after their arrival. All non-Polynesian languages of New Caledonia and the Loyalty Islands have copied the Proto Polynesian word *\*puaka* with the meaning 'pig'. Something similar seems to have happened with dogs, though over a much wider area. For example, most languages in the Temotu Province (Reefs and Santa Cruz Islands) of Solomon Islands, in northeast, central and southern Vanuatu, and in the Loyalties show a form for 'dog' derived from Proto Polynesian *\*kulii* (Lynch 1991b).

There are cases of more substantial influence of Polynesian Outlier languages on their neighbours. For example, in southern Vanuatu, the non-Polynesian languages have copied terms for parts of the canoe, sailing and fishing technology, the names of winds, and terms associated with kava-drinking from the neighbouring Polynesian Outlier Futuna-Aniwa, which has in turn copied names of moieties, terms for some varieties of breadfruit and taro and, surprisingly in view of what was said above, the word for 'pig' from these non-Polynesian languages (Lynch 1994b, 1996; Lynch and Fakamuria 1994).

#### 4.3.2 Direct and indirect inheritance in Rotuman

The classic exposition of contact between Polynesian and non-Polynesian languages is Biggs' (1965) study of Rotuman.

Rotuman words exhibit two sets of correspondences with proto-forms. Those set I and set II reflexes which differ in shape are called diagnostic. The diagnostic members of the same set may co-occur, but no diagnostic member of one set co-occurs with diagnostic members of the other set. ...

I propose to speak of directly [i.e. set I] and indirectly inherited words [i.e. set II] rather than inherited and loan words in order to emphasize that *all* of the words

TABLE 3: SOME ROTUMAN REFLEXES OF PROTO OCEANIC CONSONANTS

| Proto Eastern Oceanic | <i>*p</i> | <i>*t</i> | <i>*k</i> | <i>*q</i> | <i>*l</i> |
|-----------------------|-----------|-----------|-----------|-----------|-----------|
| Rotuman (Set I)       | h         | f         | ʔ         | ∅         | l         |
| Rotuman (Set II)      | f         | t         | k         | ʔ         | r         |

TABLE 4: DIRECTLY AND INDIRECTLY INHERITED FORMS IN ROTUMAN

| Directly inherited forms |             |      |  | Indirectly inherited forms |               |       |
|--------------------------|-------------|------|--|----------------------------|---------------|-------|
| <i>*puke</i>             | 'uncover' → | huʔe |  | <i>*paka-</i>              | 'causative' → | faka- |
| <i>*pili</i>             | 'choose' →  | hili |  | <i>*pulu</i>               | 'hair' →      | furu  |
| <i>*taqu</i>             | 'season' →  | fau  |  | <i>*toqa</i>               | 'brave' →     | toʔa  |

with etymologies were once part of a language ancestral to Rotuman in the comparativist's sense. Some of them however re-entered Rotuman from a collateral related language after undergoing changes other than those which affected forms which had remained continuously in the Rotuman line. (Biggs 1965:389–390)

Some examples of diagnostic sets of sound correspondences are shown in Table 3, while Table 4 compares some directly and indirectly inherited words. The ancestral language in this case is what Biggs calls Proto Eastern Oceanic (probably equivalent to Proto Oceanic; Ch. 4, §1).

There are in fact a number of doublets – cases where the same protoform has been inherited directly and indirectly, with slight semantic changes. For example, *\*toka* 'come ashore' has been inherited directly as /foʔa/ 'come ashore', and again indirectly as /toka/ 'settle down'.

Biggs (1965:411) points out that 'Rotuman traditions are definite in associating at least two occupations of their island with the Samoa-Tonga area', and the set II correspondences are consistent with the interpretation that one or more languages in this area have had considerable influence on Rotuman – to the extent that, of Biggs' corpus of Rotuman words with known etymologies, 38% are directly inherited and 29% indirectly inherited (with 33% being indeterminate).

Although Rotuman is the clearest and best known case of direct and indirect inheritance from different Oceanic sources, Yapese, as well as having borrowed from non-Oceanic Palauan, also has doublets resulting from direct inheritance from its putative Melanesian ancestor and indirect inheritance through Micronesian languages such as Ulithian. There are also at least two cases of direct and indirect inheritance in southeast Papua. On the mainland, Wagawaga is a language of the Suauc group which has borrowed heavily from Tawala of the Are-Taupota group. Gumawana, in the tiny Amphlett Islands, is a language of the North Mainland-D'Entrecasteaux group whose dependency relationship with speakers of the rather different Kilivila language of the Trobriands has led to so much admixture that it was first taken to be a member of the Kilivila group (Ross 1992).

#### 4.4 Contact with intrusive languages

Contact with languages not native to the Pacific has obviously been much more recent and, because of this, is generally much more superficial in nature, except in cases like Hawai'i, for example, where the dominance of English has led to the near disappearance of the Hawaiian language. In most cases, this contact has resulted in lexical copying of terms referring to newly introduced items, ideas or social and religious practices. English has been the main contributor, directly in some cases, indirectly through Melanesian Pidgin in others. However, the languages of other colonial powers – French, German, Dutch, Japanese and Spanish – have made their contributions to the lexicons of some Oceanic languages, as have (though to a much smaller extent) the languages of immigrant Indian, Chinese, Vietnamese and other Asian communities.

Many Micronesian languages, for example, show the influence of a chequered colonial history in their vocabulary. Most of western and central Micronesia was under Spanish control from the late seventeenth century until Spain lost the Spanish-American War in 1898, at which time Guam was ceded to the United States, and Germany (which had already colonised the Marshall Islands to the east) took over the rest of Spain's

possessions. Japan succeeded Germany at the outbreak of World War I, only to see Micronesia become an American Trust Territory at the end of World War II. The influence of each of the colonial languages can be seen in many Micronesian languages – as, for example, in Trukese (Goodenough and Sugita 1980):

|               |                                   |                   |                   |
|---------------|-----------------------------------|-------------------|-------------------|
|               |                                   | Trukese           |                   |
| from Spanish  | <i>anteojos</i>                   | <i>antiyos</i>    | 'fishing goggles' |
|               | <i>koopwure</i> 'corrugated iron' | <i>cobre</i>      | 'copper'          |
|               | <i>padre</i>                      | <i>paatere</i>    | 'priest'          |
| from German   | <i>Koffer</i> 'trunk'             | <i>kiiwüfer</i>   | 'suitcase'        |
|               | <i>Gummi</i>                      | <i>kkumi</i>      | 'rubber'          |
|               | <i>Mark</i> 'currency unit'       | <i>maak</i>       | 'money'           |
| from Japanese | <i>kôen</i> 'park'                | <i>kooyeng</i>    | 'playground'      |
|               | <i>oshiroi</i>                    | <i>osiroy</i>     | 'baby powder'     |
|               | <i>ramune</i>                     | <i>ramüne</i>     | 'marbles'         |
| from English  | <i>million</i>                    | <i>miniyon</i>    | 'million'         |
|               | <i>blackboard</i>                 | <i>pinakpwoot</i> | 'blackboard'      |

The emergence of Pacific Pidgin Englishes in the last century was also a result of this kind of contact. Although most of these have died out, their descendants in Melanesia – Tok Pisin in PNG, Pijin in Solomon Islands, and Bislama in Vanuatu – are flourishing languages with official or unofficial national language status in those countries.

## 5 BRIEF HISTORY OF RESEARCH

The *Current Trends in Linguistics* volume edited by Sebeok (1971) contains an outline of the history of research into Oceanic languages of all regions. Grace (1976b) covers the New Guinea area in more detail; Schütz (1972) looks at Fiji; while Tryon (1994) contains a large bibliography of works on Oceanic and non-Oceanic Austronesian languages. Some other useful bibliographies are Carrington (1996) for the New Guinea area (including Papuan languages as well), Simons (1976) for Solomon Islands, and Lynch (1994a) for Vanuatu.

### 5.1 Up until World War II

Apart from brief wordlists published by the early European navigators, and more thorough, but quite sporadic, publications by the occasional anthropologist or administrator, the bulk of the research published on Oceanic languages before the outbreak of World War II was the result, or by-product, of Christian missionary activity. Orthographies were developed – some excellent, others of more dubious quality (Ch. 2, §5). Grammar sketches and dictionaries were published, and works of a religious (and occasionally also secular) nature were printed in a number of Oceanic languages.

The languages best known in the literature up until the 1940s were those of Fiji and Polynesia – especially Standard (Bauan) Fijian, Tongan, Samoan, Hawaiian, New Zealand Māori, Rarotongan and Marquesan. Few languages elsewhere in the Pacific had a comparable coverage, though there were substantial (for the period) publications on Jabêm, Motu, Dobu and Patpatar (PNG), and on Mota and Anejoñ (Vanuatu). Mention should also be made here of the useful compendia of grammar sketches of von der Gabelentz (1861–73), Codrington (1885) and Ray (1926), though there are obviously limitations on their usability in view of the tremendous advances in linguistics in the intervening period.

Comparative work within Austronesian still focused more on the languages to the west of Oceanic, though Dempwolff's monumental three-volume study (1934, 1937, 1938) included significant amounts of Oceanic material, and Ray (1926, *inter alia*) was also to the fore in this area.

## **5.2 More recent descriptive studies**

While missionary organisations – perhaps most notably the Summer Institute of Linguistics – have still had considerable involvement in the description of Oceanic languages, the bulk of the descriptive work since the Second World War has been done by university-based linguists. The United States' assumption of administrative responsibility for the Trust Territory of the Pacific Islands (i.e. Micronesia) saw a dramatic increase in our knowledge of these languages. Initial studies were more applied, with the aim of assisting in the work of education; but these studies often had a more 'academic' side as well.

The decision in 1966 to send Peace Corps volunteers to Micronesia meant that language courses had to be written, since Peace Corps volunteers were required to learn the local language, and this provided a fresh impetus for linguistic research. These language lessons often developed into full-scale grammars and dictionaries, mainly under the auspices of the University of Hawai'i, which continues to be the major centre for the study of Micronesian languages, and which has published or sponsored major descriptive studies on Yapese, Ulithian, Woleaian, Puluwatese, Trukese, Ponapean, Mokilese, Kosraean and Marshallese.

Work has continued also in Fiji and Polynesia, though a number of Polynesian languages which could be classified as 'well-described' before the War have received little attention since then. Recent advances in linguistics mean that some of these should probably be now considered less well-described than was originally thought. On the other hand, a number of Polynesian Outliers, not well known before the War, have received the attention of linguists – Sikaiana, Luangiua, Kapingamarangi and Futuna-Aniwa, to mention just a few.

Major advances have also been made in Melanesia. The task of describing some 400 languages is a daunting one, and the situation is still very patchy, some areas being quite well known but others still neglected. Of particular importance has been the work of a number of French linguists in New Caledonia and the Loyalty Islands (Haudricourt, Rivierre, Ozanne-Rivierre and Moyse-Faurie), which has improved very significantly our knowledge of these highly aberrant languages.

## **5.3 More recent comparative studies**

The 'boom' in descriptive studies has seen a comparable dramatic increase in comparative studies of various kinds. In the last twenty years or so, there have been very considerable advances in the reconstruction of Proto Oceanic (phonology, grammar and lexicon) and of its various intermediate protolanguages, along with detailed subgrouping hypotheses within both Oceanic and the wider Austronesian family. Initially concentrating rather more on the better-known languages of eastern Oceania, these studies have increasingly incorporated the languages of western Melanesia. Mention should be made especially of the contributions to the reconstruction of Proto Oceanic or its major subgroups by linguists such as Biggs, Blust, Grace, Pawley and, more recently, Ross.

We should also note here the major linguistic surveys of Vanuatu (Tryon 1976) and Solomon Islands (Tryon and Hackman 1983), and of the linguistic mapping work resulting from these and similar surveys by scholars at the Australian National University (Wurm and Hattori 1981). These surveys have significantly improved our knowledge of the overall picture – if only to point out where the major descriptive and comparative gaps are.

#### 5.4 Current state of knowledge

In summary, we could probably say that the languages of Micronesia, Fiji and Polynesia are now pretty well described, though relatively old grammars and dictionaries remain the standard works on some of these languages. Survey work has given us a better idea of the number and nature of the Oceanic languages in Melanesia, but less than 10% of them can be called well described. This area must obviously be the focus for research over the coming decades.

In terms of comparative studies, the nature of Proto Oceanic is reasonably clear, though there are still a number of problematical areas. Major attention will need to be focused on subgrouping issues, since this is one area where there is no clear agreement at this stage.

### 6 LANGUAGE NAMES

Language names can be a problem in the Pacific. Some languages are known by one and only one name, which may be a people's own name for themselves or their language (e.g., Motu, Suau, Arosi), or an anglicised version of a local name (e.g. Rotuman, Tongan, Samoan). In other areas, however, people do not have an actual name for their own language, referring to it simply as 'the language', 'our language', or 'correct/good language'.

Many languages of this kind, however, have often been given a geographically based name by missionaries or linguists. In Vanuatu, for example, the language spoken on the island of Mota in Vanuatu is known simply as Mota, while the language spoken near Port Sandwich in Malakula is known as Port Sandwich. Even when people do have their own name for a language, some other, usually geographical, name is often more common. Thus, two Tanna languages, spoken in the Lenakel and Whitesands areas, whose 'real' names are respectively Netvaar and Nīrak, are almost universally known, by both outsiders and by younger native speakers, as Lenakel and Whitesands.

There are also many cases where the same language is known by a variety of different names – perhaps a name in the local language plus a geographical name, or a series of names for different dialects or different localities in the language-area. There are cases where named varieties refer to different dialects, but where the language itself has no local name (and where, in fact, the people do not always recognise that the differences are at the dialect level and not the language level). The Solomon Islands language (or dialect chain) known to linguists as West Guadalcanal (Tryon and Hackman 1983), for example, has a number of named dialects, some of which appear in the linguistic literature as if they were separate languages (e.g. Gari or Ghari, Kerebutu, Nggae, Sughu and Vaturanga). Early mission studies often used the name of the area where the mission was located as the name of the language. Thus, the Raga language of Pentecost in Vanuatu has been referred to as both Lamalanga and Loltong, after two important villages. Hyphenated language names (e.g. Mono-Alu in Solomon Islands)

often indicate that there are (at least) two named dialects but no overall local name for the language. And the same name may receive different spellings. A dialect of the Vangunu language of New Georgia, for example, has been variously spelled Bareke, Bariki, Mbareke and Mbariki.

The language spoken by the Tolai people of the Gazelle Peninsula of New Britain will illustrate these problems nicely:

The language of the Tolai people, which is nowadays simply called 'Tolai' in Austronesian linguistics, has been given several names. The Tolai people themselves call it **A Tinata Tuna**, literally *the indigenous language* or **Kuanua**, which is originally a word of the language of the Duke of York Islands meaning *over there* and which was first used by the Methodist missionaries who started their mission in these islands. The Catholic missionaries introduced the names **Tuna**, literally *indigenous*, **Gunantuna**; other names used by Europeans are **Blanche Bay Dialect**, **New Britain Dialect**, **Nordgazellen Sprache**, **Neu-Pommerische Sprache** and **Raluana**. (Mosel 1984:4; references omitted).

In this book, we will use the most generally accepted name for any language, and will consistently refer to the same language by the same name (except, obviously, in direct quotations)

The names of intermediate protolanguages normally indicate either the geographical area covered by members of that subgroup or, in the case of small subgroups, the name of the largest or best-known language in the group. A few examples from Ross (1988) will illustrate this: geographical names include Proto Meso-Melanesian, Proto Huon Gulf, Proto Choiseul and Proto South-West New Britain, while names of the second type include Proto Mengen, Proto Bel and Proto Madak. Until fairly recently, protolanguage names were written with a hyphen after the prefix: Proto-Oceanic. Recent convention is to dispense with this hyphen: Proto Oceanic.

In comparative studies, two abbreviatory styles for names of languages and protolanguages are fairly frequently used. One employs three-letter abbreviations, usually in capitals, and this has been in common use until recently. The other, following Reid (1992), retains capitals only where they occur in the original. A couple of examples of both styles follow:

|                    |     |     |
|--------------------|-----|-----|
| Proto Austronesian | PAN | PAn |
| Proto Oceanic      | POC | POc |
| Anejom̃            | ANJ | Anj |
| Blablanga          | BLA | Bla |
| Duke of York       | DOY | DoY |
| West Futuna        | WFU | WFu |

---

# **SOCIOLINGUISTIC BACKGROUND**

---

## **1 SOCIO-CULTURAL BACKGROUND**

Oceanic languages are spoken by people who live in a wide variety of social, political and economic circumstances. We find people living in fully independent political entities such as Tonga (which was never formally colonised), recently independent states such as Samoa and Solomon Islands, self-governing but not fully independent territories such as the Cook Islands (associated with New Zealand) and the Federated States of Micronesia (associated with the United States of America), still dependent territories of various kinds such as Tokelau (a New Zealand dependency), New Caledonia (a French dependency) and Easter Island (a Chilean dependency), as well as the politically fully integrated Māori in New Zealand, Hawaiians in Hawai'i, and Melanesians in Irian Jaya (which is now an integral part of Indonesia).

Popular perceptions of speakers of Oceanic languages often have people lying around in idyllic splendour on beautiful beaches, or enthusiastically engaging in bloodthirsty cannibal feasts. Given that there have been more than two centuries of contact with the non-Oceanic world, much has obviously changed, assuming that these stereotypes were once even partly true. The lifestyles of speakers of Oceanic languages today range from the highly urbanised cultures of the Māori and the Hawaiians, to the predominantly rural ways of life of people in most of the remaining parts of the Oceanic world. Degrees of acculturation into western lifestyles also vary enormously, though even among people leading highly westernised urban lifestyles, traditional values and cultural activities are generally still distinctive in at least some obvious respects.

Although Christianity has been adopted with enthusiasm throughout most of the Oceanic world, there are still pockets of animist adherence in some of the more remote parts of Melanesia. Even in those places where people have adopted Christianity, a distinctively local world view generally still prevails, with a variety of local spiritual belief systems, including the power of local spirits and magic, often operating in conjunction with Christian beliefs. People in many parts of Oceania still retain much of the traditional knowledge about the habits of local fauna, and the medicinal values of local flora.

Obviously, however, much has changed since European contact. The cash economy has made major inroads into people's lifestyles. Western dress is usually worn, with traditional dress typically being worn only on special occasions. The knowledge and skills that enabled Polynesian and Micronesian sailors to make some of the earliest and most daring voyages of maritime discovery in human history are rapidly being lost, and in many cases have already disappeared. Despite popular opinion in the Christian west, cannibalism was not universally practised in Oceania; where it was once practised, it no longer is. Warfare was generally only a small-scale and highly localised activity, and is now almost unknown among speakers of Oceanic languages.

## 2 LANGUAGE VARIETIES

It is probably safe to say that no language in the world – not even the smallest – has just a single code that is used in all situations, and by all speakers. Although Oceanic speech communities are small in terms of world averages, the range of cultural backgrounds is reflected in an astonishing range of different kinds of contextually determined language varieties.

In a survey such as this, it is inevitable that attention will focus on what seems to be most ‘exotic’, but even here, there is plenty to choose from. Some of the kinds of contextually determined usages that we find in Oceania include the following:

- (1) In Tongan, there is large-scale lexical replacement of many of the most commonly used words if one is speaking to the king. Tongan society is essentially feudal, and between the king and the commoners, there is an intermediate class of hereditary *hou'eiki* ‘nobles’. For each of these social groupings, there is a separate set of lexical items that is used, e.g. commoner *hā'ele* ‘come’ corresponds to the *hou'eiki* word *me'a* and to the royal word *ha'u* (Shumway 1971:603–4). Systems of lesser degrees of complexity, often involving just ‘chiefly’ vocabulary as against ‘ordinary’ vocabulary, are found in many – though by no means all – parts of Polynesia and Micronesia, where societies are generally based on hereditary chieftainships.
- (2) In the Big Nambas language of Malakula in Vanuatu, it was traditionally prohibited for a woman to say the name of a chief, her senior male in-laws, or her eldest son, or even words that sounded like any of these names. Since people’s names were often the same as everyday words, women had to make use of several strategies that would allow them to avoid this name taboo. In some cases they would use an acceptable synonym or near synonym. Thus *tau* ‘put’ would be replaced by *uln* ‘let go of’. In other cases, there was a separate set of vocabulary held in ‘reserve’ for women to use specifically to avoid name taboo. Thus if *nauei* ‘water’ were under taboo for a particular woman, she could refer to water using the special word *tarah* instead (Fox 1996).
- (3) On the island of Ngatik near Pohnpei in Micronesia, there is reportedly a special ‘men’s language’, which consists of ordinary Ngatikese vocabulary with an admixture of words that derive from contact with European sailors in the 1830s who spoke South Seas Jargon, an early form of what is now referred to as Melanesian Pidgin (Clark 1979–80:35).

Patterns of word taboo such as that reported for Big Nambas are fairly widespread. The Tahitians, for example, were also forbidden to say the name of a chief, or even part of the name of a chief, with special vocabulary for use in such situations. This kind of cultural behaviour can obviously have a major influence on the applicability of the comparative method of historical reconstruction, as it would be possible for some languages to undergo accelerated lexical replacement. This has the effect of reducing the amount of vocabulary that can be compared with other languages in reconstruction. The existence of such cultural practices also reduces the reliability of the method of lexicostatistical comparison to determine the closeness of linguistic relationships among Oceanic languages.

The use of these kinds of systems of lexical choice is contracting in many Oceanic societies today. In particular, the system of chiefly name taboo in Tahiti is no longer practised, and Big Nambas women under the age of fifty or so no longer make the same full set of lexical substitutions made by older women. While the Tongan lexically

marked chiefly style is still vigorous – with the teaching of the special vocabulary being included in the school curriculum – some younger better-educated people are able to ‘subvert’ the system by opting to speak to nobility or royalty in English instead.

Melanesian societies are generally organised along more egalitarian lines than is the case in Micronesia and Polynesia. Even so, people – almost invariably men – can gain high status by acquiring valued possessions such as pigs or shell money. Not surprisingly, these societies are generally not marked by the existence of ‘chiefly’ speech styles. Contextually determined linguistic choices in these societies tend to operate more at the level of interpersonal relations, rather than being determined by an individual’s position within a hierarchy.

In Kabana (West New Britain), people typically have personal names that refer to everyday objects. In this society, as in many other Melanesian societies, there is a strong restriction against saying the names of one’s in-laws. This is true even if someone wants to refer to the actual thing that their in-law is named after and they are not using the word as a personal name at all. Kabana-speakers use reserved items that are either words in Kabana itself, but having a different meaning, or else words with the same meanings copied from neighbouring languages. For example, the word in Kabana for a particular kind of fish is /urae/. If your in-law is called Urae, this fish must be referred to instead as /moi/, which is usually the word for ‘taro’. The word for ‘crocodile’ in Kabana is /puaea/, but this word cannot be used if your in-law is called Puaea, and the crocodile must be referred to instead as /bagele/. This form is apparently borrowed from a nearby language, where the word for ‘crocodile’ is /vaye/. Patterns such as these are described in Chowning (1985).

Another kind of restriction in the Markham Valley involves place name taboo. Among speakers of Wampar, certain places are regarded as sacred, perhaps because the people’s ancestors’ blood has been spilt there, or because their ancestors are buried there. It is believed that, if people were to use the words these place names are derived from, the ancestral spirits would punish them with disasters, sickness, or the failure of crops upon which they depend for food. The people of this area also have a similar kind of restriction to Kabana of not saying the names of in-laws (Holzknecht 1988).

People in the Markham Valley have a range of options available to them that allow them to talk about things and at the same time avoid breaking these taboos. Some languages have two or three synonymous terms to refer to the same thing, especially for very common words. Another possibility is for people to substitute a word that is semantically related to the taboo word in some way. For example, in Mari, if the word /zah/ ‘fire’ is restricted by the place-name taboo, the word /pakap/ ‘ashes’ can be used to talk about fire instead (Holzknecht 1988).

### 3 VERNACULARS AND LINGUA FRANCA

In those parts of Oceania which are linguistically the most fragmented, contact with speakers of other languages is inevitable. The historical linguist is confronted with this fact repeatedly in that it is frequently necessary to recognise the existence of structural and lexical diffusion across language or subgroup boundaries. Diffusion has even left us with languages such as Maisin in PNG and the Santa Cruz languages in Solomon Islands where there has been real debate as to whether these languages are Austronesian or non-Austronesian, as described in Ch. 1, §4.1.

It should be pointed out that it is not only in Melanesia that we find evidence for linguistic diffusion. By combining the facts of linguistic subgrouping and archaeology,

we can arrive at a fairly good idea of how, when, and in what sequence the islands of Polynesia were settled. However, it is clear that settlement did not proceed with population movements being followed by total isolation. There is plenty of linguistic evidence for continued contact between people even after linguistic diversification had taken place. Rotuman, documented in Ch. 1, §4.3, with its multiple layers of directly inherited and borrowed vocabulary, is a well-known case in point.

Phenomena such as these presuppose the existence of widespread bilingualism over long periods of time. The bilingual (or multilingual) tradition has continued into the present. In those areas that have the largest numbers of small languages, especially in Melanesia, it is common to find people today who speak two or three different Oceanic languages (and sometimes also people who speak both Oceanic and non-Austronesian languages).

The use of particular languages exclusively as in-group vernaculars and others as out-group lingua francas has probably affected the distribution of linguistic features in ways that go beyond the spread of areal features. Thurston (1987, 1989) distinguishes between exoterogeny and esoterogeny in linguistic evolution. An exoteric language is one that functions as a contact language, and which, as a result, develops in the direction of structural simplicity. An esoteric language, on the other hand, is one that functions solely as an in-group language. Such languages tend to develop in the direction of greater complexity.

Prior to European contact, structurally and lexically reduced contact languages had sometimes developed, particularly for use among people in trading relationships. Trade was frequently conducted between coast-oriented Oceanic speakers and more inland-oriented non-Austronesian speakers in PNG. However, the resulting contact languages that emerged for use on the *hiri* trading expeditions along the Gulf of Papua between speakers of the Oceanic language Motu and non-Austronesian speakers further along the coast were predominantly based on non-Austronesian lexical input (Dutton 1985:20–35).

Oral tradition and the evidence of linguistic copying indicates that some Oceanic languages were also widely used outside their original home area. However, the traditional linguistic ecology of the region has been so radically altered by patterns of language use brought about by colonial contact that we will probably never be able to reconstruct the original areas where these were used as second languages. Nor will we ever be able to know for sure which languages were at all widely used outside their home areas. We do know, however, that Tongan must have been widely used in parts of Fiji during the long pre-contact period in which Tongans held military power over much of Fiji.

Soon after the first colonial contacts, the traditional patterns of multilingualism began to change rapidly and drastically. In many areas, European missionaries were the first agents of change to enter an area. As far as possible, they aimed to learn the local language and to operate exclusively through it. In Polynesia, these missionaries generally managed to successfully learn languages such as Māori, Tongan and Samoan, and they developed writing systems and written literatures – largely of an ecclesiastical nature – in many of these languages.

In Melanesia, the policy of operating through the local language proved more difficult to promote in many communities. Arguably also, the Oceanic languages to the west – with their more complex phonologies and morphologies – were perhaps structurally more challenging to the average newly arrived missionary. A compromise solution that developed in some areas was to choose one local language and to propagate it to neighbouring areas as a lingua franca for evangelical purposes.

Oceanic languages that were spread in this way included Nakanamanga (of Nguna, but widely used on Efate and in the Shepherd Islands), Mota (in the Banks Islands, but widely used in the Banks and Torres Islands, and also in the southeastern Solomon Islands), Roviana (widely used in the western Solomon Islands), Motu (along the central Papuan coast), Suau (on the southwest tail of Papua), Dobu (in the islands off eastern Papua), Wedau (on the north coast of the tail of Papua and adjacent islands), Gedaged (in parts of Madang Province), Jabêm (in parts of Morobe Province) and Tolai (in northern New Britain, the Duke of York Islands and New Ireland).

Some Polynesian languages were also extended beyond the area of their traditional use as a result of missionary activity. Samoan was widely used in Tokelau and Tuvalu, for example, and there are many words of recent Samoan origin that have made their way into these languages through this contact. In dialectally very diverse Fiji, the variety of the chiefly island of Bau was also adopted as a *lingua franca* over the whole archipelago. As a result of this missionary choice, modern standard written Fijian represents a continuation of a nineteenth century missionary understanding of Bauan Fijian.

New *lingua francas* have developed in a number of different parts of the Oceanic world since colonial contact. Such languages include the pidginised variety of Fijian that emerged initially as a result of multilingual contact between speakers of Fijian and imported plantation labourers in the late nineteenth century (Siegel 1987:98–127). In Port Moresby in 1884, the establishment of a colonial police force saw the spread of a structurally and lexically reduced form of the local Motu language. Because of this early association, it was known for a long time as ‘Police Motu’. This language eventually spread as a *lingua franca* throughout much of the then colony of Papua, and its descendant is still widely spoken today, though it is now generally referred to as ‘Hiri Motu’ (Dutton 1985:59–81).

Pidgin Fijian is rapidly being replaced as an inter-ethnic *lingua franca* in Fiji by English, as the economy has shifted from its original plantation basis and the population becomes more educated. While Hiri Motu still claims several hundred thousand speakers, it is no longer spreading in PNG, and Tok Pisin (see below) is making some inroads into areas where Hiri Motu had been used as a *lingua franca* until recently. It is not difficult to imagine a time when Pidgin Fijian will have completely disappeared, and Hiri Motu will have become very seriously marginalised as a *lingua franca*.

Perhaps the greatest change of all to the linguistic ecology of Melanesia has been the development of the English-lexifier contact language that we refer to generically as Melanesian Pidgin. In the second half of the nineteenth century, many thousands of speakers of Oceanic languages from the Loyalty Islands, Vanuatu, Solomon Islands, New Britain and New Ireland were recruited as plantation labourers in Samoa, Fiji and coastal Queensland. These labourers rapidly ended up speaking a contact language that was lexically derived from English, and which has a structure that shows the imprint of many of the common features of the Oceanic languages of Melanesia.

This language was repatriated to Melanesia with the end of recruiting around the turn of the century, where it developed separately into what is now referred to in PNG as Tok Pisin, in Solomon Islands as Pijin, and in Vanuatu as Bislama (Mühlhäusler 1987, Keesing 1988, Crowley 1990). These three named varieties are effectively dialects of a single language since there is a high degree of mutual intelligibility, though they do differ structurally and lexically.

Although this language was also introduced into the Loyalty Islands, it was fairly rapidly replaced by French as the *lingua franca* there. In the other parts of Melanesia, however, the language continued to spread as local plantation economies and labour markets developed. It is now used by the vast majority of adults in Vanuatu and Solomon Islands, and it is spreading rapidly in PNG, to the point where well over half the population of that country uses it as their primary *lingua franca*. Of course, increasing numbers of younger people also speak English (or, in Vanuatu, French) to some extent.

As a result of the introduction of Melanesian Pidgin, the original patterns of individual multilingualism in many places have been significantly affected. Among younger uneducated Melanesians, it is common for people to speak just two languages – their local vernacular, and Melanesian Pidgin – whereas members of earlier generations would probably have spoken three or four vernaculars and no Melanesian Pidgin. Increasing numbers of younger Melanesians are now also growing up speaking Melanesian Pidgin exclusively, or with greater confidence than their parents' vernaculars.

The final element in the linguistic make-up of Oceania involves languages introduced from outside. The most widely distributed introduced language is, of course, English, which is taught in schools in most parts of the Oceanic-speaking world. In areas that have come under French colonial influence, French is the language of education, and the national language of Indonesia, Bahasa Indonesia, just impinges on the area in which Oceanic languages are spoken. Spanish has been introduced to Easter Island. Finally, German, Dutch, Spanish and Japanese were used for short periods in various parts of the Oceanic world in the late nineteenth century and in the first half of the twentieth century as a result of colonial control or military invasion.

There are also languages introduced from outside that have come in purely as immigrant languages, rather than as colonial languages. In this category, we can include varieties of Chinese spoken by established communities of immigrant traders in most Pacific towns. Indian immigrants to Fiji have evolved a distinct variety, Fiji Hindi, which is in a diglossic relationship with the structurally and lexically quite different standard Hindi (Siegel 1987:185–210), though standard Hindi is spoken in formal contexts, and other Indian languages are spoken by small groups. There are also small communities of Vietnamese speakers in Vanuatu and New Caledonia.

Not surprisingly, where there is a *lingua franca*, its use is reflected in various vernaculars in the form of lexical copying. Words of French origin have been introduced into the vernaculars of New Caledonia and other French-influenced parts of Oceania, while English-derived words are found in many other Oceanic languages. In much of Melanesia, words of apparent English origin have in fact generally been introduced in the first instance from Melanesian Pidgin rather than English, as relatively few people actually use English in Melanesia on a day-to-day basis.

The structural influence of colonial languages on Oceanic languages has generally been minimal, and we must reject Mühlhäusler's (1987:16) claim that many of these languages today are little more than relexified English. The only major exceptions are the demographically swamped Māori and Hawaiian languages, which are spoken by many younger speakers according to patterns that clearly reflect those of English.

#### 4 LANGUAGE STATUS

When linguists speak of the status of languages, or of language varieties, they often distinguish between 'prestigious' and 'stigmatised' codes. We can apply these terms also

to the situation in Oceania today, with vernaculars by and large being considered 'stigmatised' and introduced languages such as English and French being considered 'prestigious'. Certainly, a knowledge of English or French is required if one intends to get ahead in the increasingly important cash economy. An individual who knows only a vernacular and no lingua franca at all is, in this day and age, largely locked into operating within his or her own community, and has few outside employment prospects. Especially in Melanesia, the relatively few people in Oceania who speak no lingua franca at all are generally spoken of disparagingly as ignorant bush-dwellers.

However, the term 'prestige' is sometimes problematic in modern sociolinguistics. We speak, for instance, of the 'overt' prestige that influences women to make greater use of the *-ing* variant in English, while men are subject to 'covert' prestige to make greater use of the so-called stigmatised *-in'* variant. In terms of getting ahead in the cash economy, vernaculars, as well as the various indigenous lingua francas, offer people little in Oceania. At the same time, however, vernaculars are for the most part positively valued as symbols of local identity and community cohesion. Thus, being unable to speak one's ancestral language is generally considered to be regrettable, and is sometimes regarded as something shameful. Programmes have even been instituted among Māori and Hawaiians to try to regenerate these languages among younger members of the communities, who have generally grown up speaking only English (Hirsh 1987:63–105).

Language status can also be considered from a legal or constitutional perspective. In this respect, there is a major difference between the Oceanic languages of Polynesia and Micronesia on the one hand, and those of Melanesia on the other. Some of the Polynesian and Micronesian countries have some kind of constitutional provisions relating to the status of the language of that country. However, in none of these cases is the Oceanic language given higher legal status than English at the national level.

The following provision in the constitution of Kiribati is quite typical, in that despite granting constitutional status to the Kiribati language, English effectively supersedes the local language in cases of dispute:

The provisions of this constitution shall be published in a Kiribati text as well as this English text, but in the event of any inconsistency between the two texts, this English text shall prevail. (*The constitution of Kiribati*, Chapter X, Section 127)

A similar provision in the Samoan constitution even appears to be self-contradictory:

The Samoan and English texts of this constitution are equally authoritative but, in the case of difference, the English text shall prevail. (*The constitution of the independent state of Western Samoa*, Part XI, Article 112)

In some countries, there is no attempt to stipulate what language, or languages, are to be considered as official languages, or as national languages, though some countries' constitutions do indicate what languages can be used in parliamentary debate. Thus, the Samoan constitution also says the following:

All debates and discussions in the Legislative Assembly shall be conducted in the Samoan language and the English language. The Minutes and the debates of the Legislative Assembly, every bill introduced therein, every paper presented thereto, and all minutes of proceedings, minutes of evidence and reports of committees of the Assembly shall be in the Samoan language and the English language. (*The constitution of the independent state of Western Samoa*, Part V, Article 54)

The only Melanesian country to make an constitutional guarantee of protection for the local vernaculars is Vanuatu, which has the following provision:

The Republic shall protect the different local languages which are part of the national heritage, and may declare one of them as a national language. (*The constitution of the Republic of Vanuatu*, Chapter 1, Section 3(2))

However, the constitution of Vanuatu declares Bislama to be the national language, with English, French and Bislama being declared co-equal 'official' languages, and English and French being co-equal as 'principal languages of education'. In PNG and Solomon Islands, there is *de facto* recognition of English as the national language, though in PNG Tok Pisin and Hiri Motu do gain some recognition as languages of which non-citizens must demonstrate a knowledge if they are to be granted citizenship.

## 5 WRITTEN FORMS

It is almost certain that no Oceanic language developed a written form prior to European contact. One puzzling issue, however, relates to the *rongorongo* symbols that were once used by speakers of Rapanui on Easter Island. About 500 characters have been preserved on wood, but the knowledge of how to interpret them was largely lost when early missionaries destroyed most of the tablets. There is some dispute as to the origins of these pictographic symbols. Fischer (1997) argues that the script developed as an indigenous post-contact response to Easter Islanders witnessing early Spaniards writing in the late 1700s. Others, however, have argued that this represented a genuine pre-contact development (Barthel 1971).

The earliest missionary influence in the Oceanic world was in Polynesia. This was perhaps a fortunate accident of history in that these missionaries were faced with languages that possessed some of the simplest phonological systems in the world, which, for the most part, made it fairly easy to develop writing systems for them.

By and large, these nineteenth missionaries operated on the principle that each separate phoneme should be represented by a separate (and single) orthographic symbol, well before the phoneme was enunciated as a fundamental linguistic concept. There were some lapses, with the glottal stop being written as the inverted apostrophe ('), which could not be capitalised like an ordinary letter. Because it looks more like a punctuation mark than a letter, it is often not written at all by native speakers of these languages. Māori /f/ was represented by means of the digraph *wh* (as in *whare* 'house' for /fare/), in recognition of the fact that, in those dialects of Māori that the earliest missionaries came into contact with, this phoneme was pronounced as [ɸ]. The very common phonemically distinct long vowels of Polynesian languages were either not distinguished from short vowels, or they were written with a macron (e.g. *ā*, as distinct from the unmarked short vowel *a*), again often omitted by native writers.

Literacy in Polynesia took off in a big way as people converted to Christianity. In New Zealand in the first part of the nineteenth century, for example, the literacy rate among Māori in their own language was higher than that of British settlers in English. Polynesians took the possession of books as an overt sign of Christianity, as most of the books that were available were hymn books and biblical translations. The early missionaries in Polynesia apparently did the job of producing a written literature well. Even today, the Māori version of the Bible, for example, is as well regarded stylistically among Māori as is the King James version of the English Bible.

As missionaries moved west, the phonologies they were faced with became more complex and more varied. The development of the orthography of Fijian is well documented, with a number of decisions being based on the phonemic principle long before scholars of language in Europe had worked out this principle for themselves. The velar nasal was represented by the single letter *g* (as it already was in some Polynesian languages). Phonetically prenasalised stops were represented by the single letters *b* and *d*. Because /<sup>h</sup>g/ could no longer be unambiguously represented as *g*, the still uncommitted letter *q* was adopted for this purpose. The phoneme /ð/ was arbitrarily represented by *c* rather than by a digraph such as *th*.

The languages of Melanesia presented further challenges, some of which were both innovatively and successfully met. The frequent labio-velars in the languages of Vanuatu were generally written as ordinary labials with superposed tildes (e.g. *p̃*), though in the far north of the islands, where the labio-velars have a very distinct velar onset, the labio-velar stop was written as *q*. When confronted with velar fricatives, the letter *g* was sometimes used, necessitating another choice for the velar nasal, such as *ñ*. In other parts of Melanesia, unusual sounds were met with a variety of innovative solutions. In Morobe Province (PNG), for example, the glottal stop was often symbolised as *c*.

Some of the missionary writing systems were less than ideal, however. Features such as vowel length were often not marked. In many cases, velar fricatives were not distinguished from the corresponding stops. Some missionaries went slightly overboard after the successful use of single symbols to represent phonetically complex sequences in Fijian. For instance, in one early orthography on Erromango, *x* was used to represent the diphthong /au/, while *c* was used to represent /oi/. Fortunately, such excesses of orthographic zeal did not normally catch on, and more rational orthographic alternatives generally prevailed.

The diversity of phonemic systems, as well as of orthographic solutions to these systems, means that many written symbols have a wide range of interpretations in different languages. For instance, the symbol *g* represents phonemic /ŋ/ in Fijian, /g/ in Motu, /ɣ/ in Raga and /<sup>h</sup>g/ in Paamese. For this reason, in each of our sketches, we have noted particular correspondences between phonemes and orthographic symbols. In cases where traditional orthographies under-differentiate phonemic contrasts, we have made arbitrary modifications to unambiguously distinguish all phonemes.

## 6 OCEANIC LANGUAGES INTO THE FUTURE

Of the Oceanic languages that were spoken two hundred years ago, some have not survived into the present. The Moriori language of the Chathams (off New Zealand), for example, gave way to both Māori and English, and subsequently English won out there over Māori. On Erromango in Vanuatu, there were originally probably five languages, and only one of these has survived the nineteenth century epidemics, cyclones and famines as a viable language towards the end of the twentieth century. The resulting fall from an estimated original population of at least 5000 to just over 300 by the 1930s led Capell (1954:107) to predict the eventual loss of all of these languages.

Population realignments associated with resettlement by missionaries and colonial administrators, as well as epidemics in the last century, especially in Melanesia, has certainly led to other instances of language loss. On the whole, however, there has not been any major loss of Oceanic languages so far. But we can be reasonably certain that a century from now, some of the Oceanic languages of today will no longer be being actively passed on. It must be conceded that, despite the best efforts of Hawaiian

activists, the future of the Hawaiian language is bleak. Māori still has between 30,000 and 50,000 native speakers in New Zealand, but the bulk of these are of an age that will not last beyond the present generation, so this must also be considered to be a seriously threatened language.

There are clearly inequalities in the linguistic ecology in Oceania (Mühlhäusler 1987). These inequalities are probably more apparent in Melanesia than in other parts of the Pacific (Dutton 1976, Lynch 1979, Lynch and Crowley 1986, Crowley 1989), as many vernaculars have extremely poorly developed written literatures, or are not written at all. Children in schools in Melanesian countries were until recently often punished for using their vernaculars on school grounds, following policies reminiscent of nineteenth century Māori schools in New Zealand.

While some languages are probably bound to disappear in Oceania, the question that remains is how *many* of these languages are under threat. Speaking of the linguistic situation in Oceania, Dixon (1991:230) says:

The tragic saga of language extinction which has swept across Australia is likely to extend into other parts of this region during the twenty-first century. An optimistic prediction is that of these c. 1980 languages perhaps 200 will be spoken in AD 2200 (some linguists would prefer a figure of twenty or thirty).

He states that every language with fewer than 10,000 speakers is at risk of extinction in the medium term, and that languages with less than 1000 speakers are 'severely' at risk (Dixon 1991:231).

If scientific predictions about rising sea levels associated with global warming turn out to be correct, many Pacific islands, especially the low-lying atolls of Micronesia and parts of Polynesia, will become uninhabitable. Some currently viable linguistic communities may need to be relocated in larger countries, such as the US, Australia and New Zealand. If this were to happen, the languages of even entire nations such as Kiribati, Tuvalu, and many of the formerly US administered islands of Micronesia would be at risk because of the demographic swamping that would result from relocation.

However, we should be careful not to paint an overly pessimistic picture of the immediate and medium term linguistic future for many parts of Oceania. Although native speakers of many vernaculars in Oceania themselves frequently argue that their own languages have bleak futures, this is often simply because the languages are copying new words from languages such as English. In fact, most Oceanic languages seem to be holding on remarkably well, with the grammatical structures of the ordinary spoken language remaining largely unaffected by English.

This view of relatively tenacious Oceanic languages may seem to be at odds with some of the viewpoints that were expressed earlier, but Schooling (1990:124) says the following about the situation in New Caledonia, where French political and educational influence has been very strong:

The cumulative evidence of the situation among New Caledonian Melanesians strongly supports the contention ... that communities characterized by dense, multiplex social networks, have strong norm-enforcement mechanisms – that they tend to be conservative and will maintain the status quo even in the face of considerable pressure to change. This is exactly what was discovered about the rural Melanesian communities of New Caledonia.

Schooling (1990:125) reports that linguists in the 1950s were predicting the disappearance of New Caledonian vernaculars within ten years. Not only did these

languages survive, but in many respects, he points out, they appear to be in a stronger position now than they were then, despite the fact that bilingualism with French has become the norm.

The resilience of many small Oceanic languages is clearly closely related to the emblematic value that is placed on them by their speakers. It should also be kept in mind that multilingualism is not something new in Melanesia, so introduced languages have simply been added to people's linguistic repertoires.

There are some parts of Melanesia where language shift appears to be under way, though in others the vernaculars seem still to be in a fairly strong position. Areas with the most vulnerable languages include parts of the Sepik River area in PNG, though the languages are for the most part non-Austronesian rather than Oceanic. In other parts of PNG, as well as in Solomon Islands, Vanuatu and New Caledonia, vernaculars appear to be more stable.

Interestingly, those areas where Melanesian vernaculars appear to be under greatest threat are those which have non-Austronesian speaking populations. It would be ridiculous to suggest that these languages are in any sense inherently more susceptible to shift. The explanation for the greater survival of the Oceanic languages probably relates more to the fact that non-Austronesian languages generally have fewer speakers, whose needs for economic development are greater. This factor perhaps serves to push speakers of these languages to 'better' themselves by abandoning their traditional languages (Crowley 1995).

---

# TYPOLOGICAL OVERVIEW

---

Although the Oceanic languages constitute a fairly well-defined subgrouping within the larger Austronesian family, as described in Chapter 1, they do certainly not constitute a typological unity. At the same time, however, there are certain patterns and structures which tend to recur over large geographical and genetic groupings of Oceanic languages. It is the purpose of this chapter to describe those structural features that are more widely distributed among Oceanic languages. We will also describe those features that are less widely distributed, but which are nevertheless found in languages spread over particular geographical areas.

No detailed typological survey of Oceanic languages has been published, though much of the comparative literature necessarily makes typological comparisons between languages (Pawley 1972, 1973, 1981, Clark 1976, Chung 1978, Lynch 1981, Ross 1988). Lynch (1998) represents an attempt to present a typological survey of the whole group, though for an introductory readership.

The content of this chapter will allow each of the sketches in this volume to be placed in an appropriate typological context. When reading a sketch, if a feature is not described in detail, and it is indicated in this chapter as typical for Oceanic languages in general – or for the languages of that particular area – then it can be assumed that this feature is characteristic of the language in question.

We can generalise, for example, that Oceanic languages for the most part exhibit two main patterns of possessive constructions, which we can refer to as ‘direct’ and ‘indirect’ possession. Therefore, we will not describe this difference in each sketch where this distinction is made. It is only if a particular language – or group of languages – does not make this distinction, or makes it in an unusual way, that a special point will be made. To facilitate comparison, this chapter and Chapter 5, which describes Proto Oceanic, have virtually the same structure as the sketches.

## 1 PHONOLOGY

This is an area in which it is very difficult to make generalisations that are applicable to members of the entire subgroup of Oceanic languages. Although the reconstructed phoneme inventories for Proto Oceanic and a number of intermediate languages are fairly uncontroversial (see Chapter 5), the kinds of phonological changes that have taken place in different languages have been so diverse as to produce a very wide range of different sorts of phonological typologies.

Even so, languages in this subgroup are frequently phonologically less complex than those of many other linguistic groupings in the world. Syllable structures tend to approximate a simple CV type, and phoneme inventories tend to be both fairly small, and characterised by relatively few complex articulations. The major exceptions to the latter generalisation involve the presence of labio-velar or labialised velar stops, nasals and fricatives in many Melanesian and Micronesian languages, as well as the occurrence

of phonetically prenasalised voiced stops which contrast with plain oral voiceless stops in many parts of Melanesia.

Stress is generally fully predictable, falling on the penultimate syllable of a word. In languages that have contrastive stress, it usually has a fairly low functional load. Contrastive vowel length is widely distributed in the eastern parts of Oceania, though it is by no means universal. Distinctive vowel length is much less common in western Oceania. Oceanic languages are almost exclusively non-tonal, though a few languages of New Caledonia and the Huon Gulf in PNG have developed phonemically contrastive tone.

## 2 NOUNS AND NOUN PHRASES

### 2.1 Pronouns

Pronominal systems generally involve a contrast between first, second and third person, with no gender distinctions. First person pronouns almost without exception distinguish between inclusive and exclusive. There is always a distinction between singular and plural, and generally also a distinct dual series. The latter usually contains an element that is historically related in some way to the numeral 'two', though there is normally no productive synchronic compounding process involved. Some languages in Melanesia also have a series with an element reflecting 'three', and a few have one with an element reflecting 'four'. The series with 'three' is trial in some languages but paucal (expressing the meaning of 'few') in others. The series with 'four' may be synchronically either paucal or plural (but is probably never genuinely quadral).

The number of languages that do not fit within these generalisations is generally small or geographically restricted. Only Kiribati, a few varieties of Fijian, and some New Guinea languages do not mark an inclusive/exclusive distinction. Languages with only a two-way number contrast in pronouns are for the most part geographically concentrated in the New Guinea area, as well as a scattering of Micronesian languages, and Sye and Nakanamanga in Vanuatu. Separate masculine and feminine third person pronouns are reported in Kilivila (Trobriand Islands), southern New Britain and Maringe (Santa Ysabel).

There are generally several separate paradigms of pronominal forms. Four kinds of paradigm are widespread in Oceanic languages, one of free forms, the others of affixes or clitics:

- (1) Independent (i.e. free) pronouns are used in citation and function as noun phrases, i.e. as topic in topic-comment constructions, and as subject, object, possessor or prepositional object.
- (2) Possessor suffixes on bound nouns and possessive classifiers indicate the person and number of a possessor; their syntax is described in §2.7 below. In some western Melanesian languages these suffixes also occur on some prepositions and on attributive adjectives indicating the person and number of, respectively, the prepositional object and the head noun.
- (3) Subject: most languages have one or more sets of preverbal morphemes – usually clitics, but sometimes prefixes or free forms – which indicate the person and number of the subject. In Melanesian languages these are often portmanteau forms which combine with the expression of the tense/aspect/mood categories of the verb (§3.2; in the grammar sketches these are often handled under the verb phrase).

- (4) Object: the fourth kind of paradigm is less widespread than the other three, but is found quite often in Melanesia and Micronesia. It consists of a set of postverbal clitics or suffixes indicating the person and number of the object.

Many languages lack one or more of these sets, and in western Melanesia, there are languages with other sets (see the discussions of Taiof and Kairiru in this volume). There are often partial formal similarities in the shapes of free and bound pronominal forms, especially the independent and object sets, but there are seldom systematic correspondences between the various paradigms. Manam (PNG) pronominal forms are typical:

|                           | 1INC    | 1EXC     | 2             | 3             |
|---------------------------|---------|----------|---------------|---------------|
| <i>Independent</i>        |         |          |               |               |
| SG                        |         | ŋa(u)    | ʔai(ʔo)       | ŋai           |
| DL                        | ʔitaru  | ʔeru     | ʔanru         | diaru         |
| PC                        | ʔitato  | ʔeto     | ʔanto         | diato         |
| PL                        | ʔita    | ʔeʔa     | ʔaŋ ~ ʔaʔamiŋ | di            |
| <i>Possessor</i>          |         |          |               |               |
| SG                        |         | -gu      | -ŋ            | -Ø ~ -na      |
| DL                        | -daru   | -mairu   | -minru        | -diaru        |
| PC                        | -dato   | -mato    | -minto        | -diato        |
| PL                        | -da     | -ma      | -miŋ          | -dia          |
| <i>Subject (realis)</i>   |         |          |               |               |
| SG                        |         | u-       | ʔu-           | i-            |
| PL                        | ta-     | ʔi       | ʔa-           | di-           |
| <i>Subject (irrealis)</i> |         |          |               |               |
| SG                        |         | m-       | go-           | ŋa-           |
| PL                        | ta-     | ga-      | ʔama-         | da-           |
| <i>Object</i>             |         |          |               |               |
| SG                        |         | -a       | -(i)ʔo        | -Ø ~ -i ~ -a  |
| DL                        | -ʔitaru | -ʔamairu | -ʔaminru      | -diaru        |
| PC                        | -ʔitato | -ʔamato  | -ʔaminto      | -diato        |
| PL                        | -ʔita   | -ʔama    | -ʔamiŋ        | -Ø ~ -i ~ -di |

From these paradigms, it will be noted that there is a recurring element /ru/ in the dual forms and /to/ in the paucal forms, related historically (though not synchronically) to the numerals /rua/ 'two' and /toli/ 'three' respectively.

As the absence of dual and paucal forms in the Manam subject paradigms illustrates, in some languages there are differences between paradigms in the semantic distinctions that are made. In Vinmavis (Vanuatu) it is the independent paradigm and the second and third persons of the possessor paradigm that lack dual forms, and where the independent pronouns make typical non-singular inclusive/exclusive and second-/third-person distinctions, these distinctions are missing from the subject prefixes:

|                    | 1INC | 1EXC  | 2   | 3  |
|--------------------|------|-------|-----|----|
| <i>Independent</i> |      |       |     |    |
| SG                 |      | no    | gu  | i  |
| PL                 | get  | gemem | gem | ar |

|                | lINC | lEXC | 2   | 3   |
|----------------|------|------|-----|-----|
| <i>Subject</i> |      |      |     |     |
| SG             |      | nV-  | u-  | i-  |
| DL             | er-  | er-  | ar- | ar- |
| PL             | it-  | it-  | at- | at- |

## 2.2 Nouns

Nouns are often categorised in two ways in Oceanic languages. Firstly, nouns are either directly or indirectly possessed. Directly possessed nouns occur with a possessor suffix (§2.1), whilst indirectly possessed nouns are unsuffixed. This structural distinction reflects a semantic distinction between inalienable and alienable possession (§2.7).

Secondly, nouns are either personal, local or common. Personal nouns include personal proper names and, in some languages, also kin terms denoting particular individuals relative to the speaker (e.g. ‘my father’). Local nouns include institutionalised place names as well as nouns denoting places so familiar in the environment that they need no further specification (e.g. ‘home’, ‘(own) village’, ‘(own) garden’, ‘bush’, ‘beach’ etc.). This category sometimes also includes directly suffixed locative part nouns (e.g. ‘inside’, ‘above’, ‘beneath’ etc.). The class of common nouns includes all other nouns, including kin and place nouns when not used in the senses just mentioned. Common nouns often fall into a smaller non-count and a larger count subcategory.

A noun’s category membership determines its syntactic behaviour (e.g. the prepositions with which it may occur), and sometimes an accompanying article. In the languages of the north Bougainville linkage, this categorisation has become a gender system with articles marking gender and number co-occurring not only with the head noun but also with an attribute (see Taiof in this volume, §2.2).

The only other genuine gender systems, in which there is concord between constituents of a noun phrase and the head, occur in the languages of the Kilivila (Trobriand Islands) group, where a numeral classifier system (§2.4) has undergone functional expansion to become a gender system. Thus, in Kilivila we find:

|                        |                |                   |                     |                     |
|------------------------|----------------|-------------------|---------------------|---------------------|
| <i>tau</i>             | <i>m-to-na</i> | <i>to-kabitam</i> | <i>ma-bwa-na</i>    | <i>bwa-tolu kai</i> |
| man                    | this-CL-this   | CL-intelligent    | this-CL-this        | CL-three tree       |
| ‘this intelligent man’ |                |                   | ‘these three trees’ |                     |

The two systems of nominal subcategorisation – directly/indirectly possessed and personal/local/common – are independent of each other, i.e. we cannot predict whether a particular personal (or common, or local) noun will be directly or indirectly possessed. Two pseudo-categorisations also occur, according to the nature of the possessive marker that an indirectly possessed noun co-occurs with (§2.7), and according to the form of the numeral classifier with which a noun co-occurs. These are pseudo-categorisations in the sense that the same noun may occur with several different markers/classifiers, so that the latter function as closed sets of attributes rather than as markers of the noun’s category membership.

Nouns are always invariant for morphological case. Oceanic languages generally do not mark number on nouns inflectionally, so that an unmarked noun can usually be used to express both singular and plural meanings. However, there is often a hierarchy of animacy, with more highly animate nouns more likely to obligatorily distinguish singular from plural, and less animate nouns making no formal distinction between

singular and plural. It is fairly common for Polynesian languages to distinguish singular and plural forms of a few kinship terms or nouns with human reference by either some form of reduplication, or a root modification pattern involving the lengthening of one or more root vowels. Thus, in Māori, singular *tangata* ‘person’ and *tuahine* ‘(man’s) sister’ correspond to plural *tāngata* ‘people’ and *tuāhine* ‘(man’s) sisters’. Languages of central and southeast Papua also tend to have separate plural forms for kinship terms. Oceanic languages also use various phrasal number-marking strategies (§2.4).

Nouns are often productively derived from verbs (and occasionally from other roots). Generally abstract nouns have zero-derivation or a suffix (or, in some languages in western Melanesia, the infix *-in-*), whilst agents, instruments and locations are derived by means of a prefix. Reduplication also occurs, deriving abstract nouns or instruments. For example, the Tabetube (southeast Papua) verb ‘work’ is *paisewa*, from which are formed *pai-paisewa* ‘work (NOUN)’ (reduplication), *to-paisewa* ‘one who works’ and *kaba-paisewa* ‘place where one works’. In Tolai (New Britain) we find *m-in-omo* ‘a drink’ from *momo* ‘drink’, *tu-tutuk* ‘a hammer’ from *tutuk* ‘hammer (VERB)’ (reduplication), *tena-papalum* ‘one who works’ from *papalum* ‘work (VERB)’.

### 2.3 Articles and demonstratives

Many Oceanic languages have articles that precede a noun phrase. These often make a distinction between singular and plural, and between common and proper, and sometimes make a more fine-grained set of semantic contrasts than this. In Fijian, for example, the distinction between common and proper is marked by the preposed articles *na* and *o* respectively:

|           |               |          |               |
|-----------|---------------|----------|---------------|
| <i>na</i> | <i>koro</i>   | <i>o</i> | <i>Tomasi</i> |
| ART       | village       | ART      | Thomas        |
|           | ‘the village’ |          | ‘Thomas’      |

Noun phrases with generic or locative/temporal reference generally do not appear with any article.

The languages of Manus, mainland New Guinea, Micronesia and Vanuatu generally do not have articles. What was historically an article has in many of the languages of Vanuatu and some of the languages of southwest New Britain been fused with the noun root, being morphologically inseparable in all, or at least most, morphosyntactic contexts (Crowley 1985). When these fused articles are separable, it is most likely to be when a noun appears as the second member of a nominal compound. Thus, in Atchin (Vanuatu), *temets* ‘ghost’ can be compounded with *naleng* ‘dance type’ to derive *temets-leng* ‘madman’, with the historical article (POc *\*na*) being deleted in the compounded form *-leng*. In the languages of the Admiralties, New Caledonia, and part of Santa Ysabel, however, fusion has progressed so far that the fused element is never separable, and the languages of New Caledonia have subsequently developed new sets of articles, which express a fine range of semantic distinctions. While many of the languages of western Melanesia have articles, there is also a wide scattering of languages that do not.

Demonstratives are often identical to locative pro-forms, and usually make a three-way distinction based on either person (i.e. near speaker, near addressee, or near neither), or relative distance (i.e. close, intermediate distance, and distant), with the form marking intermediate distance being the least marked and often serving as the relative clause marker or occasionally as third person pronoun. More complex systems

marking distinctions such as inland/seaward or upper/lower are found in various languages, while others make only a simple proximate/distant distinction.

## 2.4 Numerals and number-marking

The most widely distributed pattern of numerals in Oceanic languages is based on a decimal system, found throughout Polynesia and Micronesia, as well as in much of Melanesia. These languages often also have separate lexical items for ‘hundred’ and ‘thousand’, and in a few cases in Micronesia, separate lexical items for 10,000, 100,000 and 1,000,000 and so on, up to 1,000,000,000.

However, in Vanuatu and New Caledonia, as well as in a wide scattering of locations further west, quinary systems are often found, with numbers higher than ‘five’ expressed as compounds based on ‘five’ or some other word. Some languages have a combination of quinary and decimal systems, with the numbers ‘six’ to ‘nine’ being compounds involving the form for ‘five’, along with a separate lexical item for ‘ten’. These languages do not generally have separate stems for ‘hundred’ or ‘thousand’.

Numerals sometimes behave like adjectives in that they are postposed to a nominal head. However, outside western Melanesia it is more common for numerals to accept some verbal inflectional morphology, while in others there is some vestigial verbal morphology.

Some languages of Micronesia and the Admiralties, as well as the languages of the Kilivila family (§2.2), are characterised by fairly elaborate systems of numeral classifiers which are either postposed to the numeral, or directly suffixed to it. There can be more than two dozen separate numeral classifiers in Micronesian languages. Note the following from Woleaian:

|               |               |                 |               |                |              |
|---------------|---------------|-----------------|---------------|----------------|--------------|
| <i>se-uw</i>  | <i>teebol</i> | <i>se-fash</i>  | <i>wa</i>     | <i>se-faiu</i> | <i>usous</i> |
| one-CL        | table         | one-CL          | canoe         | one-CL         | bead         |
| ‘one table’   |               | ‘one canoe’     |               | ‘one bead’     |              |
|               |               |                 |               |                |              |
| <i>se-mal</i> | <i>mal</i>    | <i>se-yal</i>   | <i>temaag</i> |                |              |
| one-CL        | bird          | one-CL          | cigarette     |                |              |
| ‘one bird’    |               | ‘one cigarette’ |               |                |              |

As we noted in §2.2, nouns are not generally inflected for number, but there are other number-marking strategies. In many languages, these are normally used only for more highly animate nouns. One, rather rare, strategy is affixation. For example, Sye (Vanuatu) has a plural prefix *ovn-*, and Micronesian languages tend to use suffixes to mark definite singular and plural. It is more common for plurality to be marked by adposed independent number markers, which are often identical in shape to the third person non-singular pronouns, e.g. Paamese:

|             |             |
|-------------|-------------|
| <i>ēhon</i> | <i>keil</i> |
| child       | 3PL         |
| ‘children’  |             |

In western Melanesia a plural marker sometimes occurs between the preposed article and the noun. In north Bougainville languages, in New Caledonia and in Polynesia the article itself has distinct singular and plural forms. In many languages the number of a subject or object noun phrase is indicated by the pronominal forms in the verb phrase (§2.1, §3.2), and in western Melanesia number is often marked on an attributive adjective (§2.5).

## 2.5 Adjectives and nominal modifiers

If an Oceanic language has a class of genuine adjectives at all, it is likely to be a small closed set of forms which is defined by the fact that its members are uninflected and can be used both as a predicate and attributively when they directly follow the nominal head that they modify. Stative meanings are more generally expressed by intransitive verbs. Some languages allow such verb roots to appear as postmodifiers to nouns with no inflectional marking, while other languages require such roots always to be inflected. In such cases, a stative modifier to a noun can only be expressed in the form of a relative clause. It is common in the languages of western Melanesia for adjectives to distinguish the number and person of their referent by means of 'possessor' suffixes (§2.1). Thus, in 'Ala'ala (central Papua, this volume), we find:

oveka nama-na  
 dog big-3SG  
 'big dog'

oveka nama-ta  
 dog big-3PL  
 'big dogs'

oni      eva      kau      to?oto?o-mu  
 you:SG TOPIC person short-2SG  
 'You are a short person'

## 2.6 Basic noun phrase structure

If a language has articles and/or number markers, these usually precede the head noun. This means that in languages which lack both, the head noun is either always or frequently the first item in a noun phrase. Languages in which plurality is marked by an independent 3PL pronoun vary: in some it precedes the head noun, in others it follows it. In most languages, all other noun modifiers follow the noun, but there are a few where a demonstrative or a possessor precedes.

## 2.7 Possession

The expression of possession is usually one of the more complex aspects of the grammar of Oceanic languages, and this is an area that has been widely studied from a typological and comparative perspective (Lynch 1973, Lichtenberk 1985). There is also a considerable amount of variation in the form of possessive constructions, though the semantic distinction between alienability and inalienability lies at the core of the main structural differences in all systems.

For the majority of Oceanic languages, there are two main possessive construction types, which we refer to as 'direct' and 'indirect' possession. In the case of direct possessive constructions, a possessor suffix (§2.1) is attached directly to the possessed noun, while with indirect possession, an uninflected possessed noun is either preceded or followed by an independent possessive constituent, which is itself marked with one of the possessor suffixes. Thus, in Fijian we find the following:

|           |                |           |              |             |
|-----------|----------------|-----------|--------------|-------------|
| <i>na</i> | <i>mata-qu</i> | <i>na</i> | <i>no-qu</i> | <i>vale</i> |
| ART       | eye-1SG        | ART       | POSS-1SG     | house       |
|           | 'my eye'       |           | 'my house'   |             |

With both direct and indirect possession, a possessor noun (as against a pronoun) generally follows the possessed noun, but the order is reversed in a number of western Melanesian languages, especially those with SOV clause order. In Micronesian and eastern Melanesian languages, the possessed NP is marked with what is generally referred to as the ‘construct’ suffix, or some other linking morpheme. The construct suffix sometimes coincides in shape with the third person singular pronominal suffix, but the two are frequently morphologically distinct. Thus, compare the following in Anejoŋ:

|               |                  |                      |                  |
|---------------|------------------|----------------------|------------------|
| <i>etma-k</i> | <i>etma-n</i>    | <i>etma-i</i>        | <i>natimarid</i> |
| father-1SG    | father-3SG       | father-CONST         | chief            |
| ‘my father’   | ‘his/her father’ | ‘the chief’s father’ |                  |

In some PNG languages, however, nominal and pronominal possession are not expressed by different constructions at all. Instead, the possessed NP is marked with a pronominal suffix that expresses the pronominal category of the possessor, and the possessor appears as a full NP. Thus, in Motu (where the possessor is preposed) we find:

|                      |                      |
|----------------------|----------------------|
| <i>(lau) tama-gu</i> | <i>Morea tama-na</i> |
| (1SG) father-1SG     | Morea father-3SG     |
| ‘my father’          | ‘Morea’s father’     |

The formal distinction between directly and indirectly possessed nouns generally corresponds to a semantic distinction between inalienable and alienable possession, with direct possession expressing semantic inalienability, and indirect possession expressing alienability. The detailed semantic content of inalienable nouns varies from language to language but, in general, it includes body parts, locative parts (‘inside’, ‘underneath’ etc.), kin terms, and often abstract nouns denoting things done to or said of the possessor. Alienable nouns are all other nouns in the language. Generally, body parts that are in some sense removable, and kin over whom one has authority or who one acquires through marriage, are often indirectly possessed.

Possessive systems making only the simple direct/indirect distinction are widely distributed in the languages of PNG. However, most Melanesian and Micronesian languages have systems that are more complicated than this in that they distinguish different kinds of alienable possession by means of a set of pronominally suffixed possessive constituents or ‘classifiers’. In western Melanesia, there is often a classifier that is used to express the possession of items for consumption (i.e. food and drink, and often items used in procuring them), as distinct from other items. In eastern Melanesia, we find separate classifiers for food and drink possessions, while some languages have up to half a dozen separate classifiers expressing different semantic categories of possession. Compare the following in Lenakel (Vanuatu):

|                            |                        |
|----------------------------|------------------------|
| <i>nite niki-k</i>         | <i>nikava nimwi-k</i>  |
| taro POSS:FOOD-1SG         | kava POSS:DRINK-1SG    |
| ‘my taro’                  | ‘my kava’              |
| <i>kolei ne-k</i>          | <i>nimwa taha-k</i>    |
| sweet.potatoPOSS:PLANT-1SG | house POSS:NEUTRAL-1SG |
| ‘my sweet potato’          | ‘my house’             |

In Micronesia, as well as in Mussau (to the north of New Ireland) and Iaai (New Caledonia), the number of possessive classifiers is much larger, with many kinds of alienable possession being formally distinguished. While Micronesian languages typically

group nouns semantically according to the kind of numeral classifier they are associated with, the semantic categories encoded by the systems of numeral classifiers and possessive classifiers do not coincide (nor do their forms): the numeral classifiers are generally based on physical form, while the possessive classifiers are generally based on function.

The possessive classifiers have sometimes been referred to as reflecting noun class systems. Unlike noun classes, however, association with a particular possessive classifier is often not fixed (Pawley and Sayaba 1990). This means that the same noun can appear in more than one possessive construction, depending on the relationship between the possessor and the possessed NP. Thus, we find examples such as the following in Paamese:

|                                      |                |                                         |                  |
|--------------------------------------|----------------|-----------------------------------------|------------------|
| <i>ani</i>                           | <i>ā-k</i>     | <i>ani</i>                              | <i>ema-k</i>     |
| coconut                              | POSS:FOOD-1SG  | coconut                                 | POSS:DRINK-1SG   |
| 'my coconut                          |                | 'my coconut                             |                  |
| (of which I intend to eat the flesh) |                | (of which I intend to drink the liquid) |                  |
|                                      |                |                                         |                  |
| <i>ani</i>                           | <i>esa-k</i>   | <i>ani</i>                              | <i>ona-k</i>     |
| coconut                              | POSS:PLANT-1SG | coconut                                 | POSS:NEUTRAL-1SG |
| 'my coconut                          |                | 'my coconut                             |                  |
| (which is growing on my land)        |                | (which I intend, perhaps, to sit on)    |                  |

There are languages in which nouns may shift fairly freely from the directly to the indirectly possessed category, but in others this mobility is quite constrained and inalienable nouns occur only in the direct possession construction.

Despite the widespread distribution of these different patterns of possession, there are some languages which exhibit rather different patterns. In the languages of Malakula (Vanuatu), there are often possessive pronominal suffixes only with singular possessors. Plural pronominal possession is expressed instead by independent pronouns in association with the construct suffix, in the same way as nominal possession. Thus, in Vinmavis we find the following:

|                  |                |            |                     |               |
|------------------|----------------|------------|---------------------|---------------|
| <i>netal-ung</i> | <i>netal-n</i> | <i>get</i> | <i>netal-n</i>      | <i>matoro</i> |
| leg-1SG          | leg-CONST      | 1INC:PL    | leg-CONST           | old.man       |
| 'my leg'         | 'our legs'     |            | 'the old man's leg' |               |

In Western Fijian, there is also a formal contrast between two kinds of inalienable possession, with kin terms accepting direct pronominal suffixes and body parts accepting instead pronominal prefixes, e.g.

|               |              |              |
|---------------|--------------|--------------|
| <i>o</i>      | <i>mna-m</i> | <i>m-ulu</i> |
| ART           | mother-2SG   | 2SG-head     |
| 'your mother' |              | 'your head'  |

In Melanesia, some languages express indirect possession with a preposition which links possessed and possessor. Thus in Nakanai (New Britain):

|           |            |           |             |            |            |           |           |               |           |            |                |
|-----------|------------|-----------|-------------|------------|------------|-----------|-----------|---------------|-----------|------------|----------------|
| <i>la</i> | <i>bua</i> | <i>te</i> | <i>Pasi</i> | <i>la</i>  | <i>bua</i> | <i>te</i> | <i>la</i> | <i>tahalo</i> | <i>la</i> | <i>bua</i> | <i>t-egite</i> |
| ART       | betelnut   | PREP      | Pasi        | ART        | betelnut   | PREP      | ART       | man           | ART       | betelnut   | PREP-3PL       |
| 'Pasi's   | betelnut'  |           |             | 'the man's | betelnut'  |           |           |               | 'their    | betelnut'  |                |

The Polynesian subgroup is characterised by a completely different pattern of possessive markers. These are usually referred to in the literature as the *a* and *o* forms because possessor and possessed nouns are linked by either the particle *a* or *o*, depending on the

semantic relationship between the possessed noun and the possessor. Thus, in Samoan, we find:

|                  |               |          |               |                         |              |          |           |                    |
|------------------|---------------|----------|---------------|-------------------------|--------------|----------|-----------|--------------------|
| <i>le</i>        | <i>paopao</i> | <i>o</i> | <i>Tāvita</i> | <i>le</i>               | <i>naifi</i> | <i>a</i> | <i>le</i> | <i>taule‘ale‘a</i> |
| ART              | canoe         | POSS     | Tavita        | ART                     | knife        | POSS     | ART       | young.man          |
| ‘Tavita’s canoe’ |               |          |               | ‘the young man’s knife’ |              |          |           |                    |

With pronominal possessors, there are two sets of preposed possessive classifiers which carry pronominal suffixes, which differ mainly in the fact that the vowel of the root is *a* or *o*. Thus, in Samoan:

|              |               |              |              |
|--------------|---------------|--------------|--------------|
| <i>lo-‘u</i> | <i>paopao</i> | <i>la-‘u</i> | <i>naifi</i> |
| POSS-1SG     | canoe         | POSS-1SG     | knife        |
| ‘my canoe’   |               | ‘my knife’   |              |

The semantic distinction between the two possessive constructions is often referred to in the literature as ‘dominant’ and ‘subordinate’ possession, with dominant possession being marked by *a*, and subordinate possession being marked by *o*. Forms that express possession with *a* in Polynesian languages generally correspond to forms that participate in indirect possessive constructions elsewhere, while *o* possession corresponds roughly to directly suffixed constructions. However, there is a greater amount of arbitrariness in the way that nouns in Polynesian languages are assigned to the two possessive constructions.

## 2.8 Relative clauses

Relative clauses are generally postposed to the nominal head, with the clause being preceded by a marker that often has some kind of broad subordinating function. These languages generally allow relativisation of NPs well down the universal Accessibility Hierarchy. With relativised NPs high on the hierarchy, there may be zero trace at the site of the relativised NP, though in languages in which verbs are obligatorily cross-referenced for subject or object, there will still be marking on the verb for the relativised NP. Thus, we find the following relative clauses in Paamese:

|                                     |             |             |                  |              |
|-------------------------------------|-------------|-------------|------------------|--------------|
| <i>meatin</i>                       | <i>keke</i> | <i>ēhon</i> | <i>Ø-les-i</i>   |              |
| person                              | REL         | child       | 3SG:REAL-see-3SG |              |
| ‘the person who the child saw’      |             |             |                  |              |
| <i>meatin</i>                       | <i>keke</i> | <i>ēhon</i> | <i>Ø-selūs</i>   | <i>min-i</i> |
| person                              | REL         | child       | 3SG:REAL-speak   | GOAL-3SG     |
| ‘the person who the child spoke to’ |             |             |                  |              |

With relativised NPs lower on the hierarchy, there is generally some kind of obligatory free form trace. There are many languages, however, which require that a trace be left with all relativised NPs.

## 3 VERBS AND VERB PHRASES

### 3.1 Verbal derivation and inflection

Verbs in Oceanic languages typically do not have extensive patterns of derivational morphology. The causative is generally expressed by a verbal prefix in Polynesian and

Micronesian languages, Rotuman and Fijian, as well as in a broad scattering of Melanesian languages. Fijian and the Polynesian languages, as well as many western Melanesian languages, also have a derivational reciprocal prefix. Western Melanesian languages also often have a prefix which derives an intransitive stative from a transitive.

Reduplication is almost universally used in Oceanic verbal morphology, as well as in noun derivation. It expresses a wide range of meanings, including randomness of action, repetition, and plurality of actors or patients. It is sometimes also used to derive intransitive from transitive verbs.

Some Oceanic languages exhibit patterns of root-initial segment mutation corresponding roughly to a distinction between realis and irrealis categories (Lynch 1975, Tryon 1979, Walsh 1982, Crowley 1991). Hote, Buang, Jabêm and Bukawa of the Huon Gulf area, PNG, have patterns in which verbs marked for realis categories reflect the basic form of the root, and irrealis categories are expressed with a verb root that reflects a historical nasal increment from earlier *\*na* 'irrealis marker' (which has often lost any phonemic nasal element synchronically) (Ross 1988: 370–372). Such a pattern is also found in Nāti on Malakula and Sye on Erromango (both in Vanuatu).

Most of the languages of central Vanuatu deliver us a typological surprise: they have the opposite pattern, such that it is the irrealis which uses the basic root form, and the realis which uses verb roots reflecting a nasal increment (from earlier *\*mV-*). Most of these languages have two sets of root forms which are distributed according to the morphosyntactic context, though Paamese and Southeast Ambrym have as many as four different mutated forms of the verb root.

In some languages subject and object pronominal morphemes (§2.1) are respectively prefixed and suffixed to the verb: see §3.2 below.

We frequently find some kind of formal marking for transitivity on verbs. Some languages have derivational morphemes, normally a maximum of two, each of which may transitive an intransitive verb, increasing its valency by adding an object. With the first, which generally has a shape that can be derived from the Proto Oceanic form *\*-i* (see Ch. 5, §3.1), the object is typically a patient. With the second, derived from Proto Oceanic *\*aki(ni)*, the object is an argument such as a location, a goal, an instrument or a cause, i.e. an argument which would otherwise be an oblique noun phrase. In the Oceanic literature these objects are often referred to respectively as 'close' and 'remote' objects. In Proto Oceanic, these suffixes were generally added to an intransitive root with a final consonant, like *\*taŋis* below, but in most Oceanic languages word-final consonants have been lost, with the result that when the ancient consonant is retained before a transitive affix it is interpreted as part of the suffix, as is Fijian /ð/ here:

|               |               |                   |                     |
|---------------|---------------|-------------------|---------------------|
| Proto Oceanic | <i>*taŋis</i> | <i>*taŋis-i-a</i> | <i>*taŋis-aki-a</i> |
|               | weep          | weep-TR-3SG:OBJ   | weep-TR-3SG:OBJ     |
| Fijian        | <i>taŋi</i>   | <i>taŋi-ða</i>    | <i>taŋi-ðaka</i>    |
|               | weep          | weep-TR:3SG:OBJ   | weep-TR:3SG:OBJ     |
|               | 'weep'        | 'cry for'         | 'cry because of'    |

Because there were a number of root-final consonants in Proto Oceanic, an outcome of this process is that the transitive morphemes in Fijian and many other Oceanic languages have acquired a variety of allomorphs, e.g. Fijian (in orthographic rather than phonemic form) *-ca*, *-ta*, *-ka*, *-va*, *-na* etc., and similarly *-caka*, *-taka* and so on. This in turn has resulted in the occurrence of etymologically unexpected consonants in this position, and, in Fijian at least, in their acquisition of specific meanings, so that from

Proto Oceanic \**soka* ‘stab, spear’ we do find the expected Fijian *coka-a* ‘spear it’, but expected \**coka-aka* ‘spear with it’ is replaced by *coka-taka*. We also find *coka-va* ‘dive toward it’, *coka-ta* ‘tackle him’ and *coka-raka* ‘spear it repeatedly’ (Schütz 1985:152; also Arms 1974). In other languages, however, these suffixes are often no longer productive, and in many of the languages of Melanesia there is only vestigial marking, or no formal marking for transitivity at all.

Passive constructions are only very rarely encountered in the languages of Melanesia. In some Polynesian and Micronesian languages, there are passive constructions. In Polynesian, this is generally expressed by a morpheme that is historically related to a transitive marker. In those Micronesian languages that have a passive construction, this is marked by a separate verbal affix.

### 3.2 Basic verb phrase structure

It is in the area of verbal morphology and verb phrase syntax that Oceanic languages generally exhibit the greatest complexity. Oceanic languages generally have preposed verbal morphemes, falling into two basic types, according to whether these morphemes are free or prefixed, as in Raga (this volume) and Paamese (both in Vanuatu) respectively:

|                  |            |                  |
|------------------|------------|------------------|
| <i>ra-m</i>      | <i>ban</i> | <i>au-va</i>     |
| 3PL-CONT         | go         | 3PL:REAL-go      |
| ‘they are going’ |            | ‘they are going’ |

Languages with preposed particles (like Raga) are found throughout Micronesia, Fiji and Polynesia, though there is a significant number in Melanesia as well. Languages with verbal prefixes (like Paamese) are widely distributed throughout Melanesia.

The most transparent pattern of subject and tense/aspect/mood (TAM) marking is that in which these categories are marked by separate free forms which precede an invariant verbal root, with preceding pronominal markers indicating subject (§2.1). The number of preverbal markers can be quite large, with several markers appearing in fixed sequence. There may also be a negative marker interposed between the last of the TAM markers and the verb. This is the kind of pattern that is found with Polynesian and Micronesian languages. Thus in Samoan:

|                               |           |           |            |            |          |              |
|-------------------------------|-----------|-----------|------------|------------|----------|--------------|
| <i>‘Ou</i>                    | <i>te</i> | <i>lē</i> | <i>fia</i> | <i>alu</i> | <i>i</i> | <i>Apia.</i> |
| 1SG                           | TAM       | NEG       | DESID      | go         | PREP     | Apia         |
| ‘I don’t want to go to Apia.’ |           |           |            |            |          |              |

Straddling the divide between free and prefixing systems are those in which a morphologically complex preverbal marker expresses a combination of subject and TAM categories, with some combinations of categories being expressed by portmanteau forms. Raga, illustrated above, is a language where the portmanteau form is not prefixed to the verb. In many of the languages of Melanesia, however, we find extensive patterns of portmanteau prefixation, with varying degrees of morphotactic complexity. In some languages, there are clearly recognisable orders of subject and TAM prefixes, while in other languages there are sets of subject/TAM markers that are essentially not morphologically divisible. A language of the latter type is Paamese. Thus:

|                  |                      |
|------------------|----------------------|
| <i>lo-loh</i>    | <i>lehe-loh</i>      |
| 1INC:PL:REAL-run | 1INC:PL:DIST.FUT-run |
| ‘we ran’         | ‘we will run’        |

Postverbal morphemes are often less tightly bound to the verb than preverbal morphemes; some are better analysed as enclitics than as suffixes, others as clause constituents than as verb phrase constituents.

When a pronominal object is expressed by a bound form, this invariably takes the shape of a postverbal clitic or suffix (§2.1). When there is only a partial set of object forms, these are more likely to be singular than plural forms, and third person than non-third. Languages with bound pronominal object markers are spoken in Micronesia, New Caledonia, parts of southern and central Vanuatu and in many of the languages of western Melanesia.

It is also quite common for a generic object to be incorporated into the verb phrase. In such cases, the verb is syntactically intransitive, there is no bound object marker, and the 'object' cannot be separated from the verb. This structure is most obvious in languages with clear verb phrase boundaries like Taiof (this volume) and in languages with ergative/absolute case-marking like Roviana (this volume) or Samoan. In Iaai (this volume) a verb with an incorporated object may itself be transitivised.

Another commonly occurring category of postverbal morphemes consists of directional enclitics, with at least two members meaning 'hither' and 'thither'. These are often cognate with the verbs 'come' and 'go' in other languages, and are presumably derived from earlier directional verb serialisations (cf. §3.3). For example, in Babatana (Choiseul):

|                                 |                    |                                    |                      |                |
|---------------------------------|--------------------|------------------------------------|----------------------|----------------|
| <i>Ra ko-qisu-me</i>            | <i>kavia kuda.</i> | <i>Göi</i>                         | <i>ma-zo-la</i>      | <i>Susuka.</i> |
| I 1SG:REAL-carry-hither         | some coconut       | s/he                               | 3SG:IRR-walk-thither | Susuka         |
| 'I have brought some coconuts.' |                    | 'S/he is going to Susuka village.' |                      |                |

The final element of a verb phrase is often an aspect morpheme, either enclitic or free. In Melanesia it is also common for manner adverbs to be incorporated into the verb phrase. Both features are illustrated in this Motu clause:

|                        |                     |
|------------------------|---------------------|
| <i>E-gwau-heni-gu</i>  | <i>dikadika-va.</i> |
| 3SG-scold-give-1SG     | bad-PAST:CONT       |
| 'He scolded me badly.' |                     |

### 3.3 Verb serialisation

Serial verb constructions of various types are encountered in a wide range of Oceanic languages (Crowley 1987, Early 1993, Sperlich 1993, Hamel 1993, Bradshaw 1993). These are more easily recognisable in languages that have inflectional prefixes and suffixes, as the initial verb in a serial verb construction is the one which typically carries the prefixed markers, while the final verb is the one which typically carries the suffixed markers. Thus, in Paamese, where the discontinuous negative marker *ro-/tei* surrounds the verb, we find:

|                                  |                |             |
|----------------------------------|----------------|-------------|
| <i>Ni-ro-kan</i>                 | <i>vīs-tei</i> | <i>ouh.</i> |
| 1SG:DIST.FUT-NEG1-eat            | try-NEG2       | yam         |
| 'I will not try to eat the yam.' |                |             |

Even so, series of verbal roots in languages where verbs are uninflected are widely distributed. They can be recognised as serial verb constructions by the fact that they share nominal arguments, and a single set of tense-aspect-mood and, where relevant, negative markers, as well as often having meanings that are not completely predictable from the meanings of their constituent verbs.

Serial verb constructions in Oceanic languages differ in the extent to which the verbs in question are structurally linked. Some languages allow serial verbs to independently choose objects, while other languages only allow a single set of subject and object arguments to a serial construction. In many languages, serialised verbs all have subject prefixes, whilst in others, only the first verb is prefixed. Some languages make a contrast between ‘nuclear’ serialisations, where the verbs are bound together and have only a single set of arguments (i.e. the serial construction behaves just like a single verb), and ‘core’ constructions, where the verbs remain separate words and usually share just one argument, any other argument being the subject or object of just one of the component verbs.

Despite this variety, we can identify certain semantic types of serialisation which recur in Oceanic languages. These types are all core constructions, but some have nuclear equivalents in some languages. The types are illustrated below from Paamese. Although the transitivity of the first verb may be determined by the construction, the transitivity of the second is not so determined.

- (1) Directional/Positional: the first verb expresses movement, the second the direction of that movement or the position reached as a result of that movement. There are two syntactic subtypes, depending on whether the first verb is intransitive or transitive. If it is intransitive, the moving object/person is the subject of both verbs, e.g.

*Meatin kail a-valus au-mai.*  
 person PL 3PL:REAL-row 2PL:REAL-come  
 ‘The people rowed hither.’

If it is transitive, the moving object/person is object of the first verb and subject of the second, e.g.

*Kaik ko-muasi-nau nau-vā netan.*  
 you 2SG:REAL-hit-1SG:OBJ 1SG:REAL-go down  
 ‘You hit me down.’

- (2) Sequential: the first verb expresses movement, the second the action that follows the movement. The verbs have the same subject. A purposive relationship between the actions is usually implied:

*Ki-ro-vā-tei kī-hol?*  
 2SG:IRR-NEG1-go-NEG2 2SG:IRR-dance  
 ‘Won’t you go dancing?’

- (3) Causative: the first verb is transitive, the second expresses the result of the action of the first. The object of the first verb is subject of the second:

*Kail a-muas vuas Ø-emat.*  
 they 3PL:REAL-hit pig 3SG:REAL-die  
 ‘They killed the pig by hitting it.’

- (4) Manner: the second verb expresses how the action of the first verb was performed. The verbs have the same subject:

*Kai Ø-mual Ø-suai.*  
 he 3SG:REAL-walk 3SG:REAL-disappear  
 ‘He was walking without being seen.’

- (5) Ambient: the implicit (third person singular) subject of the second verb is the sub-event expressed by the first:

|                   |            |             |                    |
|-------------------|------------|-------------|--------------------|
| <i>Ki-hulī-n</i>  | <i>ato</i> | <i>kail</i> | <i>he-mal.</i>     |
| 2SG:IRR-count-OBJ | chicken    | PL          | 3SG:IRR-be.correct |

‘Count the chickens correctly.’

The second verb is often ‘finish’, expressing completive aspect or sequence (‘and then’).

This categorisation is certainly not exhaustive: languages such as Loniū (Admiralties), Jabêm (Huon Gulf, PNG) and Paamese have a variety of other types of serialisation. But despite this, and although serial verb constructions are widely encountered, we generally do not find the same amount of freedom in Oceanic languages that we find, for example, in many of the non-Austronesian languages of PNG, which are more thoroughgoing in the extent of their serialisation patterns. What we typically find in Oceanic languages is that only some categories of verbs can appear as the second member of serial constructions other than (b), or that there is even a lexically determined set of verbs that can be serialised.

Indeed, in some languages, some serialised verbs may even never occur as independent verbs. In certain languages of central Vanuatu, verbs which occupy the second slot in a serial construction appear to be becoming restricted to that structural slot alone. In some of these languages, these forms can be better analysed as adverbial constituents within a structurally expanded verb phrase. In some languages of the Huon Gulf of PNG, this process is complete for causative serialisation, and former verbs must now be analysed as adverbs.

Other languages have what can be analysed as verbal derivational suffixes. In languages such as Paamese and Lewo in central Vanuatu, there is some ambiguity between nuclear serialisation, adverbial postmodification and derivational suffixation with some verbal constructions, which suggests that earlier serialised verbs are currently being morphologised as derivational suffixes.

There are yet other languages in which serial verbs have been further grammaticalised, having evolved into derivational affixes of various kinds. In many languages of southeast Papua and the north coast of PNG, verbs are found with what are referred to as classificatory prefixes. These prefixes are derived from the first verb in an earlier causative serial construction. Thus all verbs expressing hitting actions, or cutting actions, or actions involving the teeth or the feet, may begin with the same prefix, with the following element either occurring independently as a verb with a related meaning, or only occurring in conjunction with one or more of these classificatory prefixes. Compare the following in Tawala, which involve the classificatory prefixes *hana-* ‘action involving teeth’ and *tu-* ‘action involving feet’:

|                    |                    |                  |                  |
|--------------------|--------------------|------------------|------------------|
| <i>hana-hedali</i> | <i>hana-loloya</i> | <i>tu-hedali</i> | <i>tu-loloya</i> |
| BITE-break         | BITE-tear          | STEP-break       | STEP-tear        |

‘break with the teeth’ ‘tear with the teeth’ ‘break with the feet’ ‘tear with the feet’

Despite being widespread, serial verb constructions are not universal in Oceanic languages. The languages of southern Vanuatu have evolved a system of echo-subject prefixes which is not found anywhere else in the Oceanic subgroup. The existence of this feature coincides with a lack of serial verb constructions. Thus, in Lenakel we find:

*R-ím-vín kani m-ím-amnuumw níkava.*  
 3SG-PAST-go and ES-PAST-drink kava  
 'He went and drank kava.'

In languages that have this construction, switch-reference can be indicated by using the regular subject marker instead of the echo-subject prefix on the second verb. Thus:

*R-ím-vín kani r-ím-amnuumw níkava.*  
 3SG-PAST-go and 3SG-PAST-drink kava  
 'He<sub>1</sub> went and he<sub>2</sub> drank kava.'

## 4 CLAUSE STRUCTURE

### 4.1 Verbless clauses

Equational sentences are generally expressed by simple juxtaposition of noun phrases with no intervening verb. However, in some languages there is a copula. In certain Vanuatu languages, a copula is optional in the present and past affirmative, though with other inflectional categories it becomes obligatory. Thus, compare the following in Paamese:

|                    |              |                             |              |
|--------------------|--------------|-----------------------------|--------------|
| <i>Mail (vī)</i>   | <i>asuv.</i> | <i>Mail ro-vi-tei</i>       | <i>asuv.</i> |
| Mail 3SG:REAL:COP  | chief        | Mail 3SG:REAL:NEG1-COP-NEG2 | chief        |
| 'Mail is a chief.' |              | 'Mail is not a chief.'      |              |

### 4.2 Verbal clauses: core arguments

A wide variety of basic constituent orders is encountered in the Oceanic subgroup. The various constituent orders are distributed geographically as follows. (Note the following abbreviatory conventions: S = subject, T = topic, V = verb phrase, O = object, X = arguments other than topic):

SVO Admiralty Islands, most Markham Valley, Siasi islands, most New Britain, New Ireland, some Bougainville, most southeast Solomons, most Vanuatu, some New Caledonia, most Micronesia.

SOV central and southeast Papua, some Markham Valley, Madang coast, Wewak coast, Sarmi coast, a few Bougainville, some New Britain.

VSO New Georgia, some Santa Ysabel, much of Polynesia, Yapese.

VOS Fijian, Anejoñ (this volume), Loyalty Islands, Kiribati, many New Caledonia, Gela (this volume).

TVX many Bougainville, Choiseul, some Santa Ysabel.

Note that clauses in which both subject and object are realised as noun phrases are rare in discourse. Clauses often consist only of a verb phrase, with its clitics or affixes coreferencing subject and object.

The presence of an independent pronoun as subject or object marks contrast or focus. However, pronominal objects are more frequently expressed by independent pronouns than are subjects.

It is clear that SVO order is geographically the most widely distributed pattern, as well as being found in the genetically most diverse sample of languages. The SOV order is restricted to certain parts of PNG, and it is generally assumed that this order has arisen as

a result of contact with non-Austronesian languages which have this as the dominant constituent order (see Ch. 1, §4.1). There are two subtypes: strict SOV (e.g. Takia, Madang coast, this volume) and 'leaky' SOV, where some – usually peripheral – constituents may follow the verb (e.g. Tawala, southeast Papua). In Fijian and many Polynesian languages, VOS and VSO respectively are generally treated as the basic word orders, but there is considerable freedom in the order in which constituents can occur.

Most Oceanic languages have fairly fixed basic constituent orders, but generally allow the movement of constituents to clause-initial position in order to express topicalisation. This has become grammaticalised in the TVX order of Bougainville, Choiseul, and Santa Ysabel. If a topic is the 'framework within which the main predication holds' (Chafe 1976:50), then T in these languages is a marked topic, a phrase whose referent is introduced for the first time or is re-introduced, often after a gap (an unmarked topic, one that is present in the immediately preceding discourse, is referenced at most by a subject or object morpheme in the verb phrase).

It might be argued that these are simply VSO or VOS languages in which topicalisation occurs, but this is not an adequate characterisation, for several reasons. First, only one argument, the 'topic', is permitted before the verb phrase (although it may be preceded by a peripheral – usually temporal – argument). Second, when the subject is not T, then its exact postverbal position is unpredictable. Third, the interrogative phrase in a *wh*-question is T in these languages, whereas in most Oceanic languages *wh*-questions have the same order as declaratives. There are a number of variations on the TVX pattern. Maringe (Santa Ysabel), for example, has a special clause-final focus position.

Oceanic languages with VO order are typologically 'well-behaved' in that they are generally associated with the occurrence of prepositions rather than postpositions, postnominal adjectives and possessed–possessor order in genitive constructions. Languages with the order OV are less well-behaved in that they associate with postpositions and possessor–possessed genitive constructions, while still having postnominal adjectives. When languages of this type have inflectionally marked subject and object affixes on the verb, the subject is typically marked by a prefix, and the object by a suffix, which is what we might expect from a VO language. The OV languages of Morobe in PNG, however, are aberrant in exhibiting a mixture of typically VO and OV features, having possessor–possessed genitive constructions, but a mixture of prepositions and postpositions.

Oceanic languages are generally nominative-accusative in their formal marking of core syntactic roles, with the distinction between subject and object being marked by word order and, in some languages, also by cross-referencing on the verb. Some languages have overt marking for subject and object roles. This is more likely to be the case with languages that do not have SVO order.

However, a few Oceanic languages have ergative-absolutive patterns of marking. A few languages of Papua (Motu, Sinagoro, Maisin) have optional ergative and absolutive clitics which follow respectively a transitive subject and an object or intransitive subject. Roviana in Solomon Islands has an absolutive morpheme which precedes an object or intransitive subject. Of the Polynesian languages, those belonging to the Tongic and Samoic subgroups have ergative marking. In these languages, both ergative and absolutive NPs are marked by preposed case markers. In some of these languages, there is a kind of split ergative pattern of marking, with some verbs marking their core nominal arguments ergatively, and other verbs being associated with accusatively marked nominal arguments.

### 4.3 Verbal clauses: peripheral arguments

Non-core nominal arguments in a clause are generally marked by adposed constituents, these typically being prepositions in the case of VO languages, and postpositions in the case of OV languages. Oceanic languages do not mark peripheral functions with affixed case markers.

Most Oceanic languages have less than half a dozen genuine adpositions, with other peripheral arguments being expressed by means of complex constructions involving an adposition and a possessive noun phrase with a locative element as its head. Thus in Motu, which has postpositions (cliticised as *ai*, which contracts to *i* after *a*), we find forms such as the following:

*ruma lalo-na=i*  
house inside-3SG:POSS-POSTP  
'in the house' (lit. 'at the house's inside')

*ruma henu-na=i*  
house underside-3SG:POSS-POSTP  
'under the house' (lit. 'at the house's underside')

*ruma lata-na=i*  
house top-3SG:POSS-POSTP  
'on top of the house' (lit. 'at the house's top')

The languages of New Caledonia, however, have larger sets of prepositions, with closer to a dozen members.

In those languages which have lost the original transitive suffixes on verbs, there is often a multi-purpose oblique preposition that is used in a semantically fairly empty way to allow an NP to be used as an object of what we might want to call a pseudo-transitive verb. Thus, in Paamese:

|                    |           |               |
|--------------------|-----------|---------------|
| <i>Na-musau</i>    | <i>en</i> | <i>sauen.</i> |
| 1SG:REAL-sing      | OBL       | song          |
| 'I sang the song.' |           |               |

### 4.4 Negative clauses

The expression of negation tends to be closely related to the expression of subject and TAM categories. In languages that express these categories with free forms, the negative marker also tends to be a free form. It is generally interposed between the TAM markers and the verb, though in some Polynesian languages, negation is marked by a clause-initial marker, while in many PNG languages it is marked clause-finally. In languages like Raga, where the subject/TAM markers are morphologically bound to each other as preposed particles, the negative marker still follows these forms and precedes the verb. In languages that have extensive inflectional prefixation, negation is generally also marked with a prefix. This form is generally morphotactically separable from the subject/TAM markers, appearing between these and the verb root.

There is a recurring tendency in Oceanic languages for negation to be expressed discontinuously. Typically, the first element occupies a preverbal slot, while the second negative element appears postverbally. This pattern is found both with languages that have free form preverbal particles (e.g. Rotuman and some Polynesian languages) as

well as languages that have inflectional prefixes (e.g. Takia, Paamese, Lenakel, Vinnavis). Thus, we find the following in Paamese:

*Ni-ro-lesi-ko-tei.*  
 1SG:DIST.FUT-NEG1-see-2SG-NEG2  
 'I will not see you.'

The bipartite negative markers are generally quite different in shape from language to language, and the patterns are scattered geographically, so these clearly represent parallel innovations. There is even one language, Lewo (Vanuatu; Early 1994) in which there is tripartite negative marking, with a single preverbal constituent and two postverbal negative markers, e.g.

*Pe wii re poli.*  
 NEG1 water NEG2 NEG3  
 'There is no water.'

Some languages express negation by means of a negative verb, with the negated verb being expressed as a complement to this. Thus in Southwest Tanna (Vanuatu) the complement is a nominalisation:

|                        |                    |                    |
|------------------------|--------------------|--------------------|
| <i>k-a-s-apwah</i>     | <i>n-ivŋin-ien</i> | <i>k-a-s-ivŋin</i> |
| 3NSG-CONT-PL-not       | NOM-eat-NOM        | 3NSG-CONT-PL-eat   |
| 'They are not eating.' |                    | 'They are eating.' |

In Fijian, the negated clause is expressed as a subordinate clause:

*Au na sega ni lako mai.*  
 1SG FUT NEG SUBORD come hither  
 'I will not come.'

## 5 IMPERATIVE AND INTERROGATIVE SENTENCES

### 5.1 Imperative sentences

An imperative verb phrase often has no marking at all, or only a preverbal subject morpheme. Prohibitions usually display a different form of negation from declaratives.

### 5.2 Interrogative sentences

Polar questions are generally expressed simply through an intonation change from declaratives, or a following questioning interjection. Except in TVX languages (§4.2) content questions typically involve an interrogative marker that does not move from the structural slot of the questioned constituent. Generally, interrogatives belong in the same word class as the questioned constituent, though interrogative verbs meaning 'do what', and sometimes 'do how', are encountered. Thus, we find in Paamese:

*Ki-hiteni he-mukave?*  
 2SG:DIST.FUT-say 3SG:DIST.FUT-do.how  
 'How will you say it?'

## 6 COMPLEX SENTENCES

Oceanic languages generally do not have especially complicated systems of overt marking of subordination, and subordinate markers often perform other functions in these languages. Relative clause markers, for example, are often similar or identical in shape to demonstratives, and reason clauses are often expressed by means of a causal preposition. There is often a single subordinator that expresses a wide range of subordinating functions. It is not uncommon for clauses to be simply juxtaposed without any linking morphemes at all. The structural relationship between clauses may be shown instead by interdependence in inflectional marking between main and subordinate clauses, with the range of categories that are expressed in subordinate clauses typically being a subset of those encountered in main clauses. Conjoined sentences are generally linked by a small set of conjunctions. There is widespread use of a lexical verb meaning 'say' marking subordinate clauses to verbs of locution or perception.

# PROTO OCEANIC

## 1 THEORETICAL AND METHODOLOGICAL PRELIMINARIES

Our subject in this chapter is the reconstruction of Proto Oceanic (POc), the language ancestral to all Oceanic languages, and its relationship to the rest of the Austronesian family. One may well ask, why reconstruct protolanguages? We will mention three reasons here. The first is that reconstructing protolanguages is part and parcel of working out the subgrouping of the languages in a family and thereby of making hypotheses about their prehistory. The second is that an important component of reconstruction is vocabulary, and this can tell us a certain amount about the culture of the people who used it, as well as about subsequent changes in that culture. The third reason is that, as we saw in Chapter 3, there is considerable typological variety among Oceanic languages, and the reconstruction of morphosyntactic change is of interest for linguistic theory.

In the classical comparative method, historical linguists compare corresponding features in related languages and construct hypotheses to explain the differences among them. If, say, one language has the words *tau* ‘person’, *tama* ‘father’ and *mata* ‘eye’, whilst another has *kau*, *kama* and *maka*, and there is a regular *t/k* correspondence between the two vocabularies, then we need to determine whether the earlier consonant was *\*t* or *\*k* or something else (an asterisk indicates a reconstruction). In this case, we can be reasonably confident that it was *\*t* because (i) we can observe that across the languages of the world the sound change *t* to *k* is hugely more common than *k* to *t*, and (ii) forms with *t* in the relevant words are found right across the Austronesian family, whereas *k* occurs only in a few languages here and there.

It follows that languages with *k* for earlier *\*t* have innovated. This innovation is reflected, for example, in ‘Ala’ala (this volume) and its neighbours Gabadi, Doura, Roro, Kuni and Mekeo. The tiny part of the Oceanic family tree that includes these languages is shown in Figure 4.1, where the language ancestral to them is simply labelled ‘Proto A’. It appears that POc and Proto Central Papuan (PCP) *\*t* became Proto A *\*k* (Innovation 1). From other correspondences, it seems that POc *\*k*, *\*q* and *\*g* (PCP *\*k*, *\*ɣ* and *\*g*) all merged as Proto A *\*ʔ* (Innovation 2). Because it is more likely that the two shared innovations occurred just once, in Proto A, than independently in each language, we can infer that these languages form a subgroup, i.e. are descended from the single language here labelled Proto A.

This illustration of the comparative method is not as simple as it looks, because innovations may occur one after the other, and their ordering needs to be figured out. In the history of Mekeo, the innovation POc *\*t* > Proto A *\*k* has been followed by further innovations: Proto A *\*k* has become zero word-initially and, in East Mekeo, *ʔ* word-medially. Sometimes, the interpretation of the ordering of innovations affects subgrouping. Figure 4.2 shows that the interpretation in the previous paragraph is wrong, because the rule ordering which accounts for the Gabadi data is different from

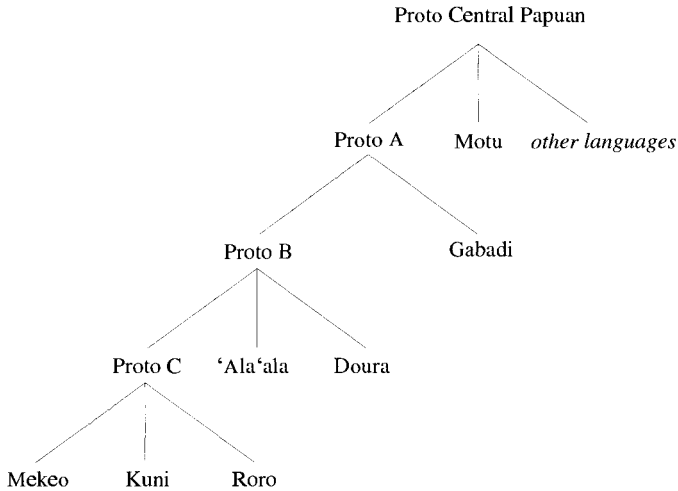


FIGURE 4.1 SUBGROUPING OF SOME OF THE CENTRAL PAPUAN LANGUAGES

the ordering which accounts for 'Ala'ala, Doura, Roro, Kuni and Mekeo. In the latter, there are three changes: PCP *\*t* became Proto B *\*k*, then PCP *\*d* became Proto B *\*t*, then PCP *\*r* became Proto B *\*d*, and they must have happened in this sequence, otherwise mergers would have occurred (and they haven't). In the case of Gabadi, however, PCP *\*d* and *\*r* first merged as Pre-Gabadi *\*d*, then the two apical stops *\*t* and *\*d* became the velar stops *\*k* and *\*g* except when they were adjacent to (before or after) the vowel *\*i*. Because change *\*t* to *k* must be ordered differently to account for the Gabadi data, the change must have happened independently in Gabadi and in Proto B, so that Innovation 1 no longer stands and only Innovation 2 supports the inference that there was a Proto A. On the other hand, the fact that the same rule ordering accounts for the 'Ala'ala, Doura, Roro, Kuni and Mekeo data allows us to hypothesise the existence of Proto B.

It may seem a little surprising to infer that the same innovation (*\*t* > *k*) has occurred in neighbouring languages independently. However, when historical linguists say that the same change has happened 'independently' in two languages, they only mean that

| POc                                         |   | Proto Central Papuan                 |   | Gabadi                                              |   | Proto Central Papuan                 |   | Proto B                              |
|---------------------------------------------|---|--------------------------------------|---|-----------------------------------------------------|---|--------------------------------------|---|--------------------------------------|
| <i>*t</i>                                   | → | <i>*t</i>                            | → | $\begin{cases} s/i \\ k \end{cases}$                |   | <i>*t</i>                            | → | <i>*k</i>                            |
| $\begin{cases} *d \\ *dr \\ *j \end{cases}$ | → | $\begin{cases} *d \\ *r \end{cases}$ |   | $\begin{cases} \text{Pre-Gabadi} \\ *d \end{cases}$ | → | $\begin{cases} d/i \\ g \end{cases}$ |   | $\begin{cases} *d \\ *t \end{cases}$ |
| $\begin{cases} *r \\ *R \end{cases}$        | → | <i>*r</i>                            |   |                                                     |   | <i>*r</i>                            | → | <i>*d</i>                            |

FIGURE 4.2 SOME SOUND CHANGES IN SOME CENTRAL PAPUAN LANGUAGES

the data in the two languages do not reflect the shared inheritance of a change which occurred in a common ancestor. They do not mean that the two changes have necessarily taken place in complete isolation. In the present case it is possible that, after Proto A had split into Pre-Gabadi and Proto B, the  $*t > k$  change took hold in one of the two languages and was adopted by speakers of the other language who interacted frequently with speakers of the first.

This small case study is abstracted from a more detailed account of the history of the Oceanic languages of Central Papua set out in Ross (1994, Forthcoming b) and illustrates the crucial role which reconstructing protolanguages plays in making hypotheses about subgrouping and linguistic prehistory. The reader who wishes to know more about the comparative method is referred to Crowley (1997), Hock (1986) or Trask (1996). Crowley (1997) uses many examples from Austronesian languages.

## 2 THE POSITION OF PROTO OCEANIC WITHIN AUSTRONESIAN

In Chapter 1, §1 (Figure 1.1), we looked briefly at the likely family tree of Austronesian. It is represented below in another format (and with an extra layer at the bottom), where each indented set of entries represents the branches into which the protolanguage above it differentiated.

Proto Austronesian

Formosan linkage (more than one subgroup?)

Proto Malayo-Polynesian (PMP)

Western Malayo-Polynesian linkage

Proto Central/Eastern Malayo-Polynesian (PCEMP)

Central Malayo-Polynesian linkage

Proto Eastern Malayo-Polynesian (PEMP)

Proto South Halmahera/Irian Jaya<sup>1</sup>

Proto Oceanic (POc)

Proto Admiralties

Western Oceanic (WOc)

Central/Eastern Oceanic (CEOc)

The reader will notice that some entries begin with 'Proto' (i.e. they are protolanguages which have parented an innovation-defined subgroup), whilst others do not (they are groups of another kind). We return to this matter in Ch. 5, §1.

The subgrouping set out above is the outcome of work by scholars using the comparative method exemplified above. The Malayo-Polynesian subgroup was proposed by Dahl (1973) and underpinned more explicitly by Blust (1977a); the Central/Eastern and Eastern Malayo-Polynesian subgroups were inferred by Blust (1978a, 1984, 1993); and Oceanic by Dempwolff (1937). In the example above, however, the hypothesised innovations were phonological. The innovations which have emerged in the process of reconstructing Austronesian linguistic prehistory are a mixture of the phonological, the morphological and the lexical (for a recent survey of work on Austronesian subgrouping, see Ross 1995b).

The status of the three groups listed below Oceanic—Admiralties, Western Oceanic and Central/Eastern Oceanic—is discussed in Chapter 5. What matters here is that by positing only three primary groupings within Oceanic, we arrive at a more secure reconstruction of POc. In the 1970s some 15–20 Oceanic groups were recognised (Grace 1968). Occasionally scholars assumed that these were primary subgroups, so that

any feature occurring in any two of these groups could be reconstructed in POc. Here we assume that POc split into only three primary subgroups, so that a POc feature can be reconstructed if it is reflected (a) in at least two of these groups, or (b) in one of these groups and in a non-Oceanic Austronesian language. A moment's thought will show that this is a more conservative approach, as it is harder to find cognate forms under these conditions than in two out of fifteen groups.

We deal here only with the reconstruction of POc, and not with lower-order Oceanic interstage languages.<sup>2</sup> A reconstructed lexicon is not provided here, as an overview has been published by Pawley and Ross (1995), reconstructed terminologies in POc and lower-order interstages are found in Pawley and Ross (1994), and a more detailed POc lexicon is in preparation (Ross, Pawley & Osmond 1998).

## 2.1 Reconstructive and notational conventions

Ideally, the reconstruction of POc that follows should be accompanied by full supporting evidence, but this lies beyond the scope of this book. Where we cite published reconstructions, we give references to these. Where there are supporting data in the sketches in this volume, we refer to them. But some of what follows is based on Ross' as yet unpublished research.

Reconstructed forms are normally marked by a preceding asterisk (\*), but asterisks are omitted from tabulations in this chapter for the sake of readability. Other conventions used in presenting these forms are:

- (*x*)    it cannot be determined whether *x* was present;
- (*x,y*)   either *x* or *y* was present, but the evidence does not tell us which;
- [*x*]    the item is reconstructible in two forms, one with and one without *x*;
- [*x,y*]   the item is reconstructible in two forms, one with *x* and one with *y*;
- x-y*   *x* and *y* are separate morphemes;
- x-*    *x* is a prefix or proclitic; or *x* is a root which takes an enclitic or a suffix
- x*    *x* is a suffix or enclitic; or *x* is a root which takes a proclitic or a prefix
- ⟨*x*⟩    *x* is an infix.

## 2.2 Proto Oceanic and Proto Malayo-Polynesian

The listing above states that the immediate ancestor of POc was PEMP, and it would seem logical to present as POc's defining innovations those which separate it from PEMP. Instead, however, POc is conventionally defined in terms of its innovations relative to PMP, partly because this was how Dempwolff originally defined them (Ch. 1, §1), and partly because neither of the intervening protolanguages, PCEMP and PEMP, has such a firm research foundation as PMP and POc. POc underwent quite significant innovations relative to PMP, and many of these probably occurred after pre-POc speakers left their kin, who presumably lived near Cenderawasih Bay in Irian Jaya, and settled in the Bismarck Archipelago, possibly on the north coast of New Britain. The very fact that these innovations can be reconstructed implies that the pre-POc/POc speech community remained fairly unified for long enough for these innovations to occur, i.e. at least for a century or two.

To facilitate comparison and reference, section 3 of this chapter has a structure similar to that of Ch. 3 and the sketches. That is, §3.1, §3.2, and §3.3 correspond respectively to §1, §2 and §3 in Ch. 3 and the sketches, and so on. In accordance with

this scheme, innovations in the POc phoneme inventory relative to PMP are set out in §3.1.1. However, certain morphosyntactic innovations between PMP and POc were system-wide and are reflected in various aspects of POc morphology, so an overview of them is given here.<sup>3</sup>

The reconstructed PMP verbal clause is typologically unusual and quite different from that of POc. Transitive verbs in PMP evidently had two basic ‘voices’,<sup>4</sup> marked by affixes on the verb. But unlike the voices of many other languages, *both* apparently had two core noun phrases. It is convenient to follow Dixon (1979) and label the referents of the two noun phrases A (actor, agent) and O. In English, for example, the basic voice (the active) is transitive and takes nominative/subject (=A) and accusative/object (=O) (e.g. *James heard a voice*), but the second voice (the passive) is intransitive and takes only a nominative (=O), A being deleted or represented by a *by*-phrase (e.g. *A voice was heard [by James]*). In some Australian languages, the basic voice is transitive and takes absolutive (=O) and ergative (=A), but the second voice (the antipassive) is intransitive and takes only an absolutive (=A), O being deleted or represented by a dative. In both language types, the forms of the core noun phrase and verb of an intransitive clause resemble those of the basic (transitive) voice.

The organisation of PMP was different from either English or Australian languages. In PMP, as in many modern Austronesian languages of the Philippines and parts of Malaysia, Indonesia, Madagascar and Taiwan, both voices were transitive, taking two core noun phrases, each introduced by a case-marking article. The basic voice (the ‘passive’) took a nominative (=O) and a genitive (=A). The genitive is so named because its case-marking was the same as for a possessor. The second voice (the ‘active’) also took a nominative (=A), but the second core noun phrase was marked as ‘accusative’. The terms are placed in inverted commas because they are used rather unconventionally. Thus the ‘passive’ is in certain respects more basic. In modern languages which retain the PMP system, the passive is the usual form for the story-line in a narrative, and the active was in a sense the ‘less transitive’ voice, as the nominative was typically definite, the ‘accusative’ typically indefinite. Unlike in the English and Australian systems, the PMP resembles the second (‘active’) voice.<sup>5</sup> These differences are summarised below:

|            | <i>Intransitive</i>    | <i>Basic voice (transitive)</i> |                 | <i>Second voice</i> |                               |
|------------|------------------------|---------------------------------|-----------------|---------------------|-------------------------------|
| English    |                        | <i>active</i>                   |                 | <i>passive</i>      |                               |
|            | nominative<br>(A-like) | nominative<br>A                 | accusative<br>O | nominative<br>O     | [ <i>by</i> -phrase=A]<br>[A] |
| Australian |                        | <i>ergative</i>                 |                 | <i>antipassive</i>  |                               |
|            | absolutive<br>(O-like) | absolutive<br>O                 | ergative<br>A   | absolutive<br>A     | [dative]<br>[O]               |
| PMP        |                        | <i>‘passive’</i>                |                 | <i>‘active’</i>     |                               |
|            | nominative<br>(A-like) | nominative<br>O                 | genitive<br>A   | nominative<br>A     | ‘accusative’<br>O             |

We illustrate the PMP system with reconstructed examples because they clarify presentation, though we certainly make no claim that they are precise in every detail. In this passive example, the verb root is \**kaRat* ‘bite’ and the passive voice marker the suffix \*-*ən*. The unmarked clause order placed the verb first, the nominative phrase last:

*\*kaRat-ən na manuk a wai*  
bite-DIRECT.PASS GNV chicken NMV mango

‘The chicken is biting the mango’ OR ‘The mango is being bitten by the chicken.’

The active voice marker was the infix *\*⟨um⟩*:

*\*k⟨um⟩aRat ta wai a manuk*  
⟨ACTIVE⟩bite ACV mango NMV chicken

‘The chicken is biting a mango.’

Some scholars would claim that PMP had four voices, rather than two, as there are three types of passive. As well as the ‘direct passive’ illustrated above, there were also a local passive formed with the suffix *\*-ən* and an instrumental or benefactive passive with the prefix *\*i-*. For example:

*\*kaʔən-an na manuk a kahiw*  
eat-LOCAL.PASS GNV chicken NMV tree

‘The chicken is eating in the tree’ OR ‘The tree is eaten in by the chicken.’

As the reconstructions show, with common noun phrases the articles were *\*a* ‘nominative’, *\*na* ‘genitive’ and *\*ta* ‘accusative’.<sup>6</sup>

The perfective was marked in all voices by the infix *\*⟨in⟩*. The direct passive suffix *\*-ən* did not co-occur with *\*⟨in⟩*, so the perfective direct passive form corresponding to *\*kaRat-ən* was *\*k⟨in⟩aRat*.

Some of the passive verb forms, at least, were also nominalisations (Ross 1995a), and retain this status in modern languages. Hence a verbless clause like the one below, where *\*k⟨in⟩aʔən* meant ‘thing eaten’, can be reconstructed, as can similar clauses with predicates *\*kaʔən-ən* ‘thing to be eaten, food’, *\*kaʔən-an* ‘eating place’ and *\*i-kaʔən* ‘eating utensil’ or ‘person eaten for’:

*\*a k⟨in⟩aʔən a wai*  
NMV ⟨PFTV:DIRECT.PASS⟩eat NMV mango  
‘The mango was what was eaten.’

The verb forms reconstructed above occurred in indicative independent clauses and as nominalisations. Other forms occurred as imperatives and in dependent clauses. The PMP morphological paradigm was as follows (where √ = root, and bolded forms are attested as nominalisations):<sup>7</sup>

| PMP                                  | nominalisation OR      |            |                         |
|--------------------------------------|------------------------|------------|-------------------------|
|                                      | independent/indicative | perfective | dependent or imperative |
| active or intransitive               | ⟨um⟩√                  | ⟨um-in⟩√   | √                       |
| active only                          | [ma]N-√                | naN-√      | paN-√                   |
| direct passive                       | √-ən                   | ⟨in⟩√      | √-a                     |
| local passive                        | √-an                   | ⟨in⟩√-an   | √-i                     |
| benefactive/<br>instrumental passive | √i-                    | i-⟨in⟩√    | √-an                    |

There was a variety of derivational morphemes, the most common of which, shown above, was the active formative *\*paN-* and its variants. *\*-N-* combined with a following obstruent to form a nasal (it surfaced as *\*-ŋ-* before a vowel). Thus the active dependent form of *\*takaw* ‘steal’ was *\*panakaw*.

Case in PMP was marked not only on articles but also on some pronoun forms. There were nominative and genitive sets, both of them enclitics to the verb, but it is probable that the two sets differed only in the singular and possible that neither set had 1EXC:PL or 2PL forms, independent pronouns being used instead:<sup>8</sup>

|            |    | 1INC        | 1EXC         | 2            | 3                                    |
|------------|----|-------------|--------------|--------------|--------------------------------------|
| nominative | SG | –           | = <i>aku</i> | = <i>kaw</i> | = <i>ya</i>                          |
| genitive   | SG | –           | = <i>ku</i>  | = <i>mu</i>  | =( <i>y</i> ) <i>a</i> , = <i>ña</i> |
| both       | PL | = <i>ta</i> | ∅            | ∅            | = <i>da</i>                          |

The reconstructions below illustrate their use:

\**k(um)aRat=kaw ta wai*  
 <ACTIVE>bite=2SG:NMV ACV mango  
 ‘You are biting a mango.’

\**kaRat-ən=mu a wai*  
 bite-DIRECT.PASS=2SG:GNV NMV mango  
 ‘You are biting the mango’ OR ‘The mango is being bitten by you.’

\**kita-ən=kaw na babinay*  
 see-DIRECT.PASS=2SG:NMV GNV woman  
 ‘The woman sees you’ OR ‘You are seen by the woman.’

If both the genitive and the nominative in a passive clause were pronouns, it appears that the genitive pronoun was cliticised to the verb and the nominative was a free form.

As well as the aspect contrast between neutral and perfective, PMP had a mood contrast between realis and irrealis, the latter formed by reduplication. It also had a number of auxiliaries which functioned as aspect markers or modals or negators, like \**qati* below. The auxiliaries preceded the verb, which took a dependent form, and attracted the actor (nominative or genitive) enclitic pronoun.

\**qati=kaw kaRat ta wai*  
 NEG=2SG:NMV bite:ACTIVE ACV mango  
 ‘You are not biting a mango.’

\**qati=mu kaRat-a a wai*  
 NEG=2SG:GNV bite-DIRECT.PASS NMV mango  
 ‘You are not biting the mango’ OR ‘The mango is not being bitten by you.’

Quite a complex series of changes to verbal clause structure occurred between the break-up of PMP and the emergence of POc. Some of the earlier changes are also reflected in a number of western Indonesian languages (Wolff 1996), but later the historical paths of these languages and of POc diverged significantly. The details of these changes lie beyond the scope of this work: what is important here is the light they cast on the morphosyntax of POc.

The first major change was that the PMP aspect/mood inflection was lost, auxiliary use increased to compensate for the losses, and as a result pronouns occurred with increasing frequency in front of the dependent verb. Eventually, this came to be the regular position for the pronoun representing the actor, and it was procliticised to the dependent verb, even when there was no auxiliary, and the dependent verb in turn was reanalysed as the main verb. Independent forms now retained only the function of nominalisation, and the ‘dependent or imperative’ column of the paradigm above became the verbal paradigm. As

a result, both nominative and genitive pronouns finished up in preverbal position, the nominatives with actives and intransitives, the genitives with passives, and this perhaps explains why two subject pronoun sets seem to be reconstructible in POc, one from nominatives, one from genitives (§3.2.1). Reconstructions of this interstage are:

\**kaw=kaRat*                      *ta*      *wai*  
2SG:NMV=bite:ACTIVE ACV    mango  
'You are biting a mango.'

\**mu=kaRat-a*                              *a*      *wai*  
2SG:GNV=bite-DIRECT.PASS NMV    mango  
'You are biting the mango' OR 'The mango is being bitten by you.'

\**kita-a=kaw*                              *na*      *babinay*  
see-DIRECT.PASS=2SG:NMV    GNV    woman  
'The woman sees you' OR 'You are seen by the woman.'

At some point, the direct passive dependent  $\sqrt{-a}$  apparently lost its suffix, rendering it identical to the active  $\sqrt{-}$ , and the direct passive function was taken over by the local passive  $\sqrt{-i}$ . Nominalisations were also affected: reflexes of  $\sqrt{-\text{en}}$  'direct passive nominaliser' crop up only in a few fossils like POc *\*kanog(a)* 'flesh, meat, coconut flesh' (Ross 1996c: 174), and the local passive nominalisers  $\sqrt{-an}$  and  $\sqrt{-in}$  (as a nominaliser the latter was perhaps an Oceanic innovation) had by POc times come to overlap functionally with  $\sqrt{-in}$  'direct passive nominaliser' (§3.2.2; Ross 1998d).

Two changes affected the benefactive/instrumental markers. The nominalisation  $\sqrt{-in}$  disappeared, and the benefactive/instrumental verb form  $\sqrt{-\text{an}}$  was replaced by  $\sqrt{-ak\text{en}}$  (this change is reflected across a large part of Indonesia). All of this resulted in a much reduced morphological paradigm:

|                                      | Nominalisation                         |                      | Verb                   |
|--------------------------------------|----------------------------------------|----------------------|------------------------|
|                                      | <i>neutral</i>                         | <i>perfective</i>    |                        |
| active or intransitive               | —                                      | —                    | $\sqrt{-}$             |
| active only                          | —                                      | —                    | $[\text{pa}]N\sqrt{-}$ |
| direct/local passive                 | $\sqrt{-an}$ , ( $\sqrt{-\text{en}}$ ) | $\sqrt{-in}$ $[-an]$ | $\sqrt{-i}$            |
| benefactive/<br>instrumental passive | $i\sqrt{-}$                            | —                    | $\sqrt{-ak\text{en}}$  |

Finally, the active voice disappeared, and with it the voice system, leaving the unaffixed root to serve only as the intransitive. This left the erstwhile passive  $\sqrt{-i}$  (which carried the heaviest functional load in a system of the PMP type) as the default transitive verb form, and the reflex of  $\sqrt{-ak\text{en}}$  as the benefactive/instrumental transitive (the 'remote transitive' reconstructed by Pawley 1973). Genitive proclitic pronouns and the article *\*na* now marked transitive subjects, and nominative proclitic pronouns and the article *\*a* marked intransitive subjects and transitive objects. At some stage—perhaps in association with the loss of the voice system—the clitic pronouns came to function as coreferencing (agreement) markers. That is, a verb had a proclitic coreferencing the subject and, if it was transitive, an enclitic coreferencing the object, regardless of whether subject and object noun phrases also occurred in the clause or not:

\**mu=kaRat-i=a*                               $[\text{ikoe}]$  *a*      *wai*  
2SG:SUBJ-bite-TR-3SG:OBJ [2SG] OBJ    mango  
'You are biting the mango.'

\**ñā=kaRat-i=a*                      *na*   *manuk*   *a*   *wai*  
 3SG:SUBJ=bite-TR=3SG:OBJ   SUBJ   bird   OBJ   mango  
 'The chicken is biting the mango.'

\**ko=kaRat*                      [*ikoe*]  
 2SG:SUBJ=bite:ITR   [2SG]  
 'You are biting.'

The reinterpretation of the root form as intransitive only was probably associated with the fact that the 'active' was in any case the less transitive voice, the 'accusative' being typically indefinite. POC retains a few relics of the PMP active/passive contrast, for example \**paŋan* 'eat' (from PMP \**paŋaʔən* < \**paN-kaʔən*) versus \**kani* 'eat' (from PMP \**kaʔən-i*). Their reflexes reveal no apparent difference in meaning, but at some stage, perhaps immediately pre-POC, there was a contrast between the following (glossed in both the PMP and the POC manner, as it not clear what the system was at the time):

\**ko=paŋan*                      [*ikoe*]*ta*                      *wai*  
 2SG:NMV=eat:ACTIVE   [2SG] ACV   mango  
 2SG:SUBJ=eat:ITR                      [2SG] INDEF   mango  
 'You are eating a mango.'

\**mu=kani=a*                      [*ikoe*]*a*                      *wai*  
 2SG:GNV=eat:PASS=3SG:NMV [2SG] NMV   mango  
 2SG:SUBJ-eat:TR=3SG:OBJ                      [2SG] OBJ   mango  
 'You are biting the mango.'

By the time POC broke up, this arrangement had changed. The article \**ta* marked indefinites, but now did so regardless of whether they were subject or object. An O argument only occurred with an intransitive if it was incorporated into the verb phrase and no morpheme intervened between verb and noun, e.g. \**paŋan wai* 'eat mango, be mango-eating'. For most verbs the intransitive/transitive contrast was simply between intransitive root (e.g. \**kaRat* 'bite:ITR') and root + transitive \*-*i* (e.g. \**kaRat-i* 'bite-TR'). and the now redundant functional distinction between the transitive and intransitive sets of subject proclitics (ex-genitive and ex-nominative) and between the two articles, \**na* 'transitive subject' (ex-genitive) and \**a* 'transitive object/intransitive subject' (ex-nominative) was in the process of being lost when POC broke up. The POC morphological paradigm by now was as follows:

| <i>POC</i>                              | <i>Nominalisation</i> |                              | <i>Verb</i>                                 |
|-----------------------------------------|-----------------------|------------------------------|---------------------------------------------|
|                                         | <i>neutral</i>        | <i>perfective</i>            |                                             |
| intransitive                            | —                     | —                            | √                                           |
| (relic transitives)                     | (√- <i>on</i> )       | —                            | ( <i>paN</i> -√, <i>N</i> -√)               |
| direct transitive                       | √- <i>an</i>          | ⟨ <i>in</i> ⟩ [ <i>-an</i> ] | √- <i>i</i>                                 |
| benefactive/<br>instrumental transitive | <i>i</i> -√           | —                            | √- <i>aki</i> ( <i>n</i> ), √- <i>akini</i> |

The expected POC reflex of \*√-*akən* is \*√-*akon*, but instead we find \*√-*aki*(*n*) and \*√-*akini* (§3.3.1). Presumably the now transitive suffix \*-*i* played a role in this irregular change, but how this happened is not clear. A number of well distributed WOC languages reflect a transitive suffix \*-*ni*, but we do not know whether this contributed to the formation of \*√-*akini* or is a reduction of it (or neither).

The developments we have surveyed here have their outcomes at a number of points in the following account of reconstructed POc morphology: in the pronominal system (§3.2.1), in the derivational morphology of nouns (§3.2.2), in the article system (§3.2.3), in the derivational and inflectional morphology of verbs (§3.3.1) and in the structure of verbal clauses (§3.4.2).

### 3 A PROTO OCEANIC GRAMMAR SKETCH

The lexical morpheme classes of POc were:

- nouns (subclasses: human, common, local; §3.2.2)
- verbs (subclasses: dynamic and stative, the latter divisible into adjectival and non-adjectival; §3.3.1)
- temporals (§3.4.3)
- locatives (§3.4.3).

There were no adjectives as such, just adjectival verbs (a subclass of stative verb) and a small class of adjectival nouns (a subclass of common noun). There may also have been a small class of underived adverbs.

The major phrase types of POc were:

- human noun phrases (subtypes: noun phrase with human noun as head, independent personal pronoun [§3.2.1])
- common noun phrases (§3.2.6; rarely replaced by pronouns, more commonly ellipted or replaced by demonstrative)
- locative phrases (subtypes: prepositional phrase with local noun as head of governed noun phrase [§3.4.3], locative pro-form [§3.2.3])
- temporal phrases (subtypes: prepositional phrase with local noun as head of governed noun phrase [§3.4.3], temporal [§3.4.3])
- adverbial phrases (subtypes: verbal preposition phrase [§3.4.3]; derived or underived adverb)
- verb phrases (§3.3.2).

Noun phrases occurred as subjects, all phrase-types as predicates.

#### 3.1 Phonology

##### 3.1.1 Phonemes

The reconstructed consonant phonemes of POc are:

|                |   |    |   |   |   |
|----------------|---|----|---|---|---|
| p <sup>w</sup> | p | t  | c | k | q |
| b <sup>w</sup> | b | d  | j | g |   |
|                |   | s  |   |   |   |
| m <sup>w</sup> | m | n  | ñ | ŋ |   |
|                |   | r  |   |   | R |
|                |   | dr |   |   |   |
|                |   | l  |   |   |   |
| w              |   |    | y |   |   |

The consonant inventory reconstructed by Dempwolff (1937) has been modified in various ways by Haudricourt (1951), Milke (1968), Grace (1969), Wolff (1974),

Lichtenberk (1978), Blust (1978b) and Ross (1988, 1989, 1996e). The inventory and orthography here are from Ross (1988), with the addition of *\*p<sup>w</sup>* (for which Ross 1996 provides supporting data). Evidence for these reconstructions, in the form of sound correspondences and supporting data, is also found in Tryon (1976), Lynch (1978a), Tryon and Hackman (1983), and, if we take Proto Eastern Oceanic as equivalent to POc, Biggs (1965), Cashmore (1969), Pawley (1972), Geraghty (1983, 1986, 1990).

Although the reconstructed inventory is fairly secure, questions remain about the phonetic values of some segments. The phonemes *\*p<sup>w</sup>*, *\*b<sup>w</sup>* and *\*m<sup>w</sup>* are known in the literature as 'labio-velars'. This orthography reflects their pronunciation in the majority of Oceanic languages in which they remain distinct, but there is evidence to suggest that they did indeed have the double articulations [k̠p], [g̠b] and [ŋ̠m] that 'labio-velar' suggests, since some languages (e.g. Mwotlap, this volume) retain these realisations, whilst others (on Malaita and in Fiji) have velar reflexes. Among the apicals, it is possible that *\*t* was dental, the others alveolar, as in a number of west Indonesian languages (Ozanne-Rivierre 1992) and in Oceanic Banoni (this volume). The voiced obstruents in the second row were also prenasalised. The phoneme *\*r* was an alveolar trill, whilst *\*dr* was a prenasalised alveolar trill, reflected thus in languages in the Admiralties and Fiji. The phoneme *\*c* is assumed to have been a voiceless palatal obstruent, because this is the articulation one would predict on the basis of non-Oceanic cognates and of its position in the inventory. However, it is distinctively reflected only in some Admiralties languages, where its reflexes are mostly alveolar liquids ([l], [r]) or glottals ([ʔ], [h]). In WOc and CEOc it has merged with *\*s*. The phoneme *\*j* is more widely reflected, as [tʃ], [dʒ] or [d], and was most likely a voiced palatal obstruent. Of the two postvelars (see Ross 1988: 31–32), *\*q* was probably a glottal stop, but its uvular stop reflexes in some languages give room for doubt, whilst *\*R* was probably a uvular trill, which is frequently lost or merged with a liquid (*\*r* or *\*l*) in daughter languages.

Superficially the POc consonant inventory appears to have had no fricatives except *\*s*, but, on the basis of widespread reflexes, it is likely that *\*[ɸ, β]*, *\*[x, ɣ]*, and *\*[z]* occurred as allophones of *\*p*, *\*k* and *\*s*.

We noted above that POc evidently underwent a set of innovations relative to PMP, and it is the reflexes of these innovations in Oceanic languages that define the Oceanic subgroup. A number of these innovations involve the consonants, as we see when we tabulate the correspondences between the reconstructed consonant inventories of PMP and POc (for discussion of the PMP consonant inventory, see Ross 1992):

|                |             |                      |          |             |             |          |             |          |          |          |          |
|----------------|-------------|----------------------|----------|-------------|-------------|----------|-------------|----------|----------|----------|----------|
| PMP            | <i>p, b</i> | –                    | <i>t</i> | <i>d, r</i> | <i>s, Z</i> | <i>j</i> | <i>k, g</i> |          |          |          |          |
| POc oral grade | <i>p</i>    | <i>p<sup>w</sup></i> | <i>t</i> | <i>r</i>    | <i>s</i>    | <i>c</i> | <i>k</i>    |          |          |          |          |
| nasal grade    | <i>b</i>    | <i>b<sup>w</sup></i> | <i>d</i> | <i>dr</i>   |             | <i>j</i> | <i>g</i>    |          |          |          |          |
| PMP            | <i>m</i>    | –                    | <i>n</i> | <i>ñ</i>    | <i>ŋ</i>    | <i>w</i> | <i>y</i>    | <i>l</i> | <i>q</i> | <i>h</i> | <i>R</i> |
| POc            | <i>m</i>    | <i>m<sup>w</sup></i> | <i>n</i> | <i>ñ</i>    | <i>ŋ</i>    | <i>w</i> | <i>y</i>    | <i>l</i> | <i>q</i> | <i>θ</i> | <i>R</i> |

The innovations which occurred over the pre-POc period were mergers and splits, the introduction of new phonemes, and one deletion, as follows:

- The PMP voiced/voiceless pairs *\*p*, *\*b* and *\*k*, *\*g* merged respectively as POc *\*p* and *\*k*. Ozanne-Rivierre (1992) suggests that the corresponding *\*t*, *\*d* merger was hindered by their putative mismatch in point of articulation (dental vs alveolar).
- The PMP pairs *\*s*, *\*Z* and *\*d*, *\*r* merged respectively as pre-POc *\*s* and *\*r*.
- Homorganic nasal + obstruent sequences, i.e. *\*mp*, *\*nt*, *\*ŋk*, *\*nr*, *\*ns* and *\*ñc* occurred in pre-POc (they are not tabulated above). Where these sequences

occurred word-medially, they were usually inherited from PMP. Where they occurred word-initially they were the outcome of a pre-POc innovation which is unpredictable and whose cause(s) unknown. These sequences became the unitary POc prenasalised voiced obstruents *\*b*, *\*d*, *\*g*, *\*dr*, and *\*j* (pre-POc nasal + *\*s* and nasal + *\*c* merged as POc *\*j*). It is possible that POc *\*d* never occurred word-initially. In the terminology introduced by Grace (1969) and widely used by Oceanic linguists, the prenasalised reflexes of the pre-POc obstruents are known as their ‘nasal grade’, the plain, unprenasalised reflexes as their ‘oral grade’.

- (d) Contrastive labio-velars *\*pʷ*, *\*bʷ* and *\*mʷ* developed. Most of the words containing a labio-velar lack non-Oceanic cognates, and it seems that the words were borrowed into POc from neighbouring Papuan languages. For example, it can be argued that *\*mʷapo(q)* ‘tarō’ was borrowed by POc speakers as they acquired taro-growing techniques from Papuan speakers (Ross 1996e). A few of these items were inherited into POc, and the labio-velar was the reflex of a labial occurring next to a round vowel. However, it is not clear in these items that the labio-velar actually occurred in POc (Blust 1981). Thus a number of Oceanic languages reflect *\*tamʷata* ‘man, husband’, derived from *\*tau* ‘body, person’ + *\*mataq* ‘unripe, immature, young’.

- (e) PMP *\*h* was deleted in POc.

The combined effect of (a) and (c) is that the PMP pair *\*p*, *\*b* merged as pre-POc *\*p*, then split into POc *\*p* and *\*b*. That is, either PMP *\*p* or PMP *\*b* could become either POc *\*p* or POc *\*b*. For example:

| PMP             |               | POc                          |
|-----------------|---------------|------------------------------|
| <i>*panas</i>   | ‘hot, warm’   | <i>*panas</i>                |
| <i>*punay</i>   | ‘wild pigeon’ | <i>*bune</i>                 |
| <i>*baqeRuh</i> | ‘new’         | <i>*paqoRu</i>               |
| <i>*beRek</i>   | ‘pig’         | <i>*boRok</i> ‘domestic pig’ |

Parallel processes affected PMP *\*k* and *\*g*, which could become either POc *\*k* or POc *\*g*. For example:

| PMP           |               | POc          |
|---------------|---------------|--------------|
| <i>*kuden</i> | ‘cooking pot’ | <i>*kuro</i> |
| <i>*kabut</i> | ‘mist’        | <i>*gabu</i> |

PMP etyma with an unambiguous initial *\*g*- are rare, but we find:

| PMP            |                 | POc                                   |
|----------------|-----------------|---------------------------------------|
| <i>*gapgap</i> | ‘stammer’       | <i>*kaka(p)</i>                       |
| <i>*gemel</i>  | ‘grasp in hand’ | <i>*gomo(l)</i> (Blust, in progress). |

The reconstructed vowel phonemes of POc are:

|          |          |
|----------|----------|
| <i>i</i> | <i>u</i> |
| <i>e</i> | <i>o</i> |
| <i>a</i> |          |

These are the outcomes of certain innovations relative to PMP, set out below:

| PMP | <i>i</i> , -uy(-) | <i>e</i> , -aw | -ay      | <i>a</i> | <i>u</i> |
|-----|-------------------|----------------|----------|----------|----------|
| POc | <i>i</i>          | <i>o</i>       | <i>e</i> | <i>a</i> | <i>u</i> |

PMP *\*e* became POc *\*o*, and the PMP word-final diphthongs *\*-uy(-)*,<sup>9</sup> *\*-aw* and *\*-ay* were simplified to POc *\*-i*, *\*-o* and *\*-e* respectively, the first two thereby merging with plain vowels.

### 3.1.2 Orthography

Two orthographies are in use in the literature for POc. The first was introduced by Biggs (1969) for Proto Eastern Oceanic, adopted for POc by Grace (1969), and has been used with a number of variants (separated by a slash) shown below. The parenthesised symbols were introduced by Blust (1978b).

The second orthography, used in this chapter, was introduced by Ross (1988: Ch. 3). The main reasons for its introduction were (i) to bring it closer to the phoneme inventories of Oceanic languages and (ii) to give emphasis to a reinterpretation of the nasal grade of the palatal obstruents. Where Grace had *\*ns* as the nasal grade corresponding to oral grade *\*s*, and *\*nj* corresponding to no oral grade phoneme, Ross showed that *\*ns* had not existed and that the nasal grade corresponding to *\*s* was *\*nj*, which he rewrote as *\*j*. He also showed that this *\*j* was the nasal grade corresponding to Blust's (1978) *\*j* in the old orthography. That is, the nasal grades of *\*s* and Blust's *\*j* had merged in POc.

|            |             |                      |                      |              |              |           |           |              |           |          |          |
|------------|-------------|----------------------|----------------------|--------------|--------------|-----------|-----------|--------------|-----------|----------|----------|
| Grace etc. | oral grade  | <i>p</i>             | <i>pw</i>            | <i>t</i>     | <i>d/r</i>   | <i>s</i>  | —         | ( <i>j</i> ) | <i>k</i>  |          |          |
|            | nasal grade | <i>mp</i>            | <i>mp/mpw</i>        | <i>nt</i>    | <i>nd/nr</i> | <i>ns</i> | <i>nj</i> | —            | <i>ŋk</i> |          |          |
| Ross       | oral grade  | <i>p</i>             | <i>p<sup>w</sup></i> | <i>t</i>     | <i>r</i>     | <i>s</i>  |           | <i>c</i>     | <i>k</i>  |          |          |
|            | nasal grade | <i>b</i>             | <i>b<sup>w</sup></i> | <i>d</i>     | <i>dr</i>    |           | <i>j</i>  |              | <i>g</i>  |          |          |
| Grace etc. | <i>m</i>    | <i>gm/mw</i>         | <i>n</i>             | ( <i>ñ</i> ) | <i>ŋ</i>     | <i>w</i>  | <i>y</i>  | <i>l</i>     | <i>q</i>  | <i>θ</i> | <i>R</i> |
| Ross       | <i>m</i>    | <i>m<sup>w</sup></i> | <i>n</i>             | <i>ñ</i>     | <i>ŋ</i>     | <i>w</i>  | <i>y</i>  | <i>l</i>     | <i>q</i>  | <i>θ</i> | <i>R</i> |

Both orthographies represent vowels in the same way.

### 3.1.3 Phonotactics

POc words were made up of (C)V syllables, with the option of a word-final consonant. These word-final consonants are lost in the majority of Oceanic languages, but regularly retained in a scattering of Western Oceanic languages, in Mussau, and in some cases in Southern Vanuatu and New Caledonian languages. They are also retained sporadically in a scattering of other languages. Quite often, as with POc *\*kaka(p)* 'stammer' and *\*gomo(l)* 'grasp in hand' above, we know that the PMP form had a final consonant, but no reflex occurs in any of the Oceanic languages which reflect final consonants, and so we have no means of knowing whether that consonant occurred in POc or not. In such cases the final consonant is shown in parentheses in the reconstructed POc form. as it appears that PMP word-final consonants were quite consistently retained in POc.

PMP permitted word-medial CVC syllables, as in *\*gapgap* 'stammer'. One of the innovations which defines POc is the loss of the final consonant of a word-medial syllable, as in POc *\*kaka(p)*. The most common context for this innovation is reduplicated forms like *\*gapgap* (Blust 1977b), but it also occurred elsewhere; for example, PMP *\*beRgi* 'night' became POc *\*boŋi*.

POc vowel sequences have to our knowledge never been systematically investigated, but they seem not to have been particularly common. A check of several geographically and genetically well distributed languages which are otherwise phonologically

conservative<sup>10</sup> reveals a consistency which probably reflects the POc pattern, namely that each vowel in a sequence is the nucleus of a separate syllable. Although some Oceanic languages contrast long vowels with short or contrast a sequence of two identical vowels with a single vowel, this kind of contrast is not reconstructed for POc, where only sequences of unlike vowels were permitted. We saw above that POc *\*e* is derived exclusively from PMP word-final *\*-ay*, and this historical origin apparently precludes its occurrence in POc vowel sequences except in a few probable borrowings and some derived forms. However, it is probable that all sequences of *\*i*, *\*a*, *\*o* and *\*u* occurred. Well attested, for example, are *\*waiR* 'fresh water', *\*raun* 'leaf', *\*maosak* 'ready to be eaten (because ripe or cooked)', *\*bou* 'main bearers supporting raised floor or roof structure, centre post supporting ridgepole', *\*panua* 'inhabited territory; community together with its land and things on it', *\*qio(r;R)* 'spear, arrow'. It is probable, incidentally, that the falling sequences *\*ua* and *\*ia* were not distinct from *\*uwa* and *\*iya*.

### 3.1.4 Stress

POc stress also remains uninvestigated, but phonologically conservative languages generally agree in displaying primary stress on the penultimate syllable and secondary stress on every second syllable preceding the penultimate, and this was probably the basic POc pattern. Manam and Fijian agree on a pattern whereby in normal speech, if the vowel of the penultimate syllable is immediately preceded by another vowel, then stress falls on that vowel, i.e. on the antepenultimate syllable. Since these two languages are both phonologically conservative and geographically and genetically far apart, this may also have been a POc pattern. If so, then we should reconstruct, for example, *\*máosak* 'ready to be eaten' rather than *\*maósak*.

## 3.2 Nouns and noun phrases

### 3.2.1 Pronouns

POc pronouns were probably organised according to the pattern described in Chapter 3, §2.1, but the forms of the subject proclitics are problematic and are discussed below. It is probable that the subject proclitic and object enclitic paradigms were defective, the gaps being filled by the corresponding independent forms.

|                                                     | 1INC   | 1EXC           | 2             | 3              |
|-----------------------------------------------------|--------|----------------|---------------|----------------|
| <i>Independent</i> (Ross 1988: 367)                 |        |                |               |                |
| SG                                                  | –      | [i]au          | [i]ko[e]      | ia             |
| NON-SG                                              | kita   | ka[m]i, kamami | ka[m]u, kamiu | [k]ira         |
| <i>Subject proclitic</i> (see the discussion below) |        |                |               |                |
| SG                                                  | –      | ku=, au=       | mu=, ko=      | (y)a=, ña=, i= |
| NON-SG                                              | ∅, ta= | ∅              | ∅             | ∅, ra=         |
| <i>Object enclitic</i> (Evans 1995)                 |        |                |               |                |
| SG                                                  | –      | =au            | =ko           | =a             |
| NON-SG                                              | ∅      | ∅              | ∅             | =ra            |
| <i>Possessor suffix</i> (Ross 1988: 112)            |        |                |               |                |
| SG                                                  | –      | -gu            | -mu           | -ña            |
| NON-SG                                              | -da    | -ma[m]i        | -m[i]u        | -dra           |

Evans (1995) has shown that the paradigm of object enclitics was defective (the gaps were filled by independent pronouns), as was the PMP nominative paradigm from which they are descended (§2.2). We do not know why this defectiveness should have remained stable over time. Certainly, it later broke down in the many Oceanic languages which have created clitic forms to fill the gaps in the paradigm.

Formally, the independent pronouns fall into two sets: (i) those which reflect the PMP nominative paradigm but are prefixed, obligatorily or optionally, with *\*i*, reflecting the PMP personal article, and (ii) those which begin with *\*k[a]-*, apparently a marker of courtesy in PMP or earlier. The POc form of the 3PL was probably *\*ira*, i.e. an *\*i*-initial form, to which *\*k-* was optionally added by analogy with the other non-singular forms.

The possessor suffixes are well attested. They show only a limited formal correspondence to the independent and object forms because they reflect the PMP long genitive forms (as opposed to the short genitive forms in §2.2), incorporating the ancient genitive elements *\*-m-* and *\*-n-* and subsequent phonological changes. For example, the 1SG independent and object forms *\*[i]au* and *\*=au* both reflect the Proto Austronesian base *\*akú*, whilst possessor *\*-gu* reflects Proto Austronesian *\*-n-akú* (with loss of unstressed *\*-a-*). The early history of the various second person forms is explained by Blust (1977a).

The reconstruction of POc subject proclitics is more complex.<sup>11</sup> Although subject proclitics (or prefixes) occur in many well distributed Oceanic languages and we can infer their presence in POc, their forms vary considerably and a number of competing reconstructions can be made. We organise these into three sets:

|                | 1INC           | 1EXC               | 2                   | 3                 |
|----------------|----------------|--------------------|---------------------|-------------------|
| <b>Set I</b>   |                |                    |                     |                   |
| SG             | —              | <i>au=</i>         | <i>ko=</i>          | <i>i=</i>         |
| NON-SG         | <i>ta=</i>     | <i>θ (?)</i>       | <i>θ (?)</i>        | <i>ra=</i>        |
| <b>Set II</b>  |                |                    |                     |                   |
| SG             | —              | <i>ku=</i>         | <i>mu=</i>          | <i>(y)a=, ña=</i> |
| NON-SG         | <i>ta= (?)</i> | <i>θ (?)</i>       | <i>θ (?)</i>        | <i>ra= (?)</i>    |
| <b>Set III</b> |                |                    |                     |                   |
| SG             | —              | <i>[y]a=</i>       | <i>o-</i>           | <i>e-</i>         |
| NON-SG         | —              | <i>ka[i]=, mi=</i> | <i>kau=, m[i]u=</i> | —                 |

From the developments reconstructed in §2.2, we see that Sets I and II respectively reflect the PMP nominative and genitive clitics, and that the identity of their plural forms is of PMP antiquity. On the basis of this reconstruction, we would expect Set I to be the intransitive subject set, Set II the transitive. However, although both sets of forms are reflected in Oceanic languages, the functional distinction between them is found nowhere, and we infer that it was being lost when POc broke up. The forms in Set I are widely reflected. Of the Set II forms, *\*ku=* 1SG is well attested in Oceanic languages. The 2SG form *\*mu=* is reflected in only one language, Yapese, but the latter is a significant witness as it is evidently a very early offshoot of Oceanic (see Ch. 5, §2). The 3SG form *\*ya=* appears to be well supported, and *\*ña=* has possible reflexes in Yapese and various Central/Eastern Oceanic languages.

The Set III forms are reduced versions of the independent pronouns, the latter sometimes being reduced in different ways in different languages. For example, the independent *\*kami* 1EXC:PL, which sometimes occurred as *\*kai*, has seemingly given rise to the subject forms *kai=*, *ka=* and *mi=* in various languages. (Alternatively, *kai=* and

*mi*= may reflect PMP nominative and genitive forms whose reconstruction is insecure.) If they are indeed reductions of independent pronouns, then the forms in Set III are probably not reconstructible for POc, but the later result of parallel innovations in different languages whereby a form derived from Set I or II was reduced by phonological attrition and a new form created by the cliticisation and subsequent reduction of the independent form (e.g. *\*ya*= 1SG:SUBJ is a reduction of *\*iau* 1SG:OBJ).

There is good evidence that the numerals *\*rua* '2', *\*tolu* '3' and perhaps *\*vat[i]* '4' were cliticised to independent and possessor non-singular forms to mark dual, trial and paucal number respectively, giving, for example, *\*/k[ira]=tolu* 'they three' and *\*=dra=tolu* 'of them three'. When they served as clitics, *\*rua* and *\*tolu* were optionally reduced to *\*ru* and *\*tou* (the latter reflected in Yapese, the Admiralties, the Willaumez languages, Fiji and Polynesia).

### 3.2.2 Nouns

POc nouns were categorised in the two ways described in Ch. 3, §2.2. Firstly, they were either directly or indirectly possessed (§3.2.7), a structural distinction which corresponded closely to the semantic distinction between inalienable and alienable possession. Whilst most, and perhaps all, nouns belonged by default to either the directly possessed or the indirectly possessed category, there were morphological devices which allowed at least some nouns to cross the boundary between the categories if possessive semantics required it. Indirectly possessed nouns occurred with one of three (or more) possessive classifiers which specified more narrowly the nature of the possessive relation (§3.2.7.2). Many of these nouns could occur with one of two or more classifiers, depending on the nature of the possessive relation.

Cutting across this categorisation was a second which divided nouns into categories which were reflected in the articles (§3.2.3) and prepositions (§3.4.3) with which they co-occurred:

- (1) Personal: personal proper names and kin terms used of known individuals,
- (2) Common:
  - (a) human: human nouns which were not in the personal category, and some non-human animates (evidently those considered more human-like, e.g. pets);
  - (b) non-human: all other common nouns
- (3) Local: proper place names, nouns denoting familiar places (e.g. 'home', '(own) village', '(own) garden', 'bush', 'beach' etc.), directly possessed locative part nouns (e.g. 'inside', 'above', 'beneath' etc.), temporal nouns.

The subcategory labels 'human' and 'non-human' are somewhat loosely applied, as the definitions above show.

Kin nouns evidently fell into both the personal and the common human categories. If they referred to identifiable individuals and were equivalent to a proper name, e.g. 'my father', they were treated as personal nouns; otherwise they were common human. Local nouns (called 'locative bases' by Pawley 1972: 33) were never preceded by an article. Local nouns other than proper place names could also function as common nouns. For example, *\*Rumaq* 'house' functioned as a local noun in the prepositional phrase *\*i Rumaq* 'at home' (§3.4.3) but as a common noun in *\*a na-gu Rumaq* 'my house' (§3.2.7.2).

Because within each of these categorisations certain nouns could belong to more than one category, neither categorisation represented a gender system.

Only some human nouns were (optionally) marked for number, probably by reduplication. The form of this reduplication varies from language to language, but in scattered Western Oceanic languages it takes the unusual form of reduplication of the stressed syllable (Ross 1998a). Given that change from unusual (marked) to commonplace (unmarked) is attested more frequently than the reverse, it seems likely that this form of reduplication (rather than the more commonplace reduplication of the initial syllable) reflects the POc process. This would allow the reconstruction of pairs like *\*papine* ‘woman’ vs *\*papipine* ‘women’, *\*tam<sup>w</sup>aqāne* ‘man’ vs *\*tam<sup>w</sup>aqaqāne* ‘men’, and perhaps *\*tamá-dra* ‘their father’ vs *\*tamamá-dra* ‘their fathers’.

POc had three apparently productive affixes, inherited from PMP, which derived nouns from verbs and sometimes from other bases. They were:

| Affix  | Example                                                  | Functions                        |
|--------|----------------------------------------------------------|----------------------------------|
| *i-√   | <i>*asa(q)</i> ‘grate:ITR’, <i>*i-asa(q)</i> ‘grater’    | instrument nominaliser           |
| *⟨in⟩√ | (see below)                                              | patient and general nominaliser  |
| *√-an  | <i>*mate</i> ‘die’, <i>*mate-an</i> ‘deathbed, cemetery’ | location and general nominaliser |

Perhaps because of the loss of PMP *\*√-ən*, which had formed nouns like *\*kaʔən-ən* ‘thing to be eaten, food’ (§2.2), there was evidently some functional overlap between POc *\*⟨in⟩√* and *\*√-an*. POc *\*⟨in⟩√* is reflected as the patient and general nominaliser only in Mussau and the Meso-Melanesian cluster. In Roviana (this volume) the general nominaliser is *⟨in⟩*, e.g. Roviana *kainera* ‘song’ from *kera* ‘sing’, whilst *\*-an* retains its local meaning, e.g. Roviana *huhuve-ana* ‘bathing place, bath’ from *huhuve* ‘bathe’ (Roviana *-ana* is the regular reflex of *\*-an*). In other, widely distributed, languages (including some in the Meso-Melanesian cluster) the reflex of the erstwhile locational *\*√-an* has taken over the function of general nominaliser, so that some reflexes of POc *\*mate-an*, e.g. Vitu (this volume) *mate-a*, Longgu (this volume) *mae-a-*, Māori *mate-ŋa*, mean ‘death’, rather than ‘deathbed’ or ‘cemetery’. It seems that this functional extension was already occurring in POc. The affix combination *\*⟨in⟩√-an* is reflected in fossilised reflexes of POc *\*kani-an* ‘food’ in the languages of Epi (central Vanuatu) (Tryon 1976:289), but it is not clear whether this occurred in POc or is a local innovation.

The history of *\*-an* also displays phonological complications: we find forms that reflect not only *\*-an* but also *\*-aŋ*, *\*-ana*, *\*-aŋa* and *\*-ŋa*. Since many languages have lost POc final *\*-a* and many more have lost POc final consonants, reconstructing the history of these forms is very difficult. This is an area where more detailed research is needed.

It seems likely that a noun-deriving innovation which arose through compounding was already productive by the break-up of POc. This was the combination of the prefixes *\*tau-* (from *\*tau* ‘person’) and *\*ka[i]-* (from *\*kaiu* ‘tree’) with a verb (or perhaps sometimes a noun) to form agentive and instrumental nouns respectively. The latter is attested in Mussau and in Papuan Tip and CEOc languages. The use of *\*kaiu* ‘tree’ to form instruments was presumably a case of grammaticisation via metonymy: *\*kaiu* was also used for ‘stick’, and a stick served numerous instrumental functions in Oceanic societies.

### 3.2.3 Articles and demonstratives

The first element of the POc noun phrase was often an article, and, like many Oceanic languages, POc apparently used one article with personal nouns and another with common non-human nouns (and no article at all with common human or with locative nouns).

A number of CEOc languages reflect an article *\*i* with personal nominals (Pawley 1972, Ross 1998c). Certain WOc languages reflect *\*i* with personal possessors, and *\*e* elsewhere. The latter is perhaps also reflected as a fossil prefix in some Anejoñ (this volume) kin terms: *etpo*- ‘grandparent’ (POc *\*tubu*-), *etma*- ‘father’ (POc *\*tama*-), *etwa*- ‘same-sex sibling’ (POc *\*tuqa*- ‘older same-sex sibling’). It is reasonable to infer that the forms reflecting *\*i* and *\*e* have a common origin and that perhaps *\*e* was the POc phrase-initial form, *\*i* phrase-internal. There is also evidence of another POc personal article *\*qa*, with reflexes in southern New Britain, Southeast Solomonic and Polynesian. How this contrasted with *\*e*/*\*i* is unknown.

Crowley (1985) reconstructs a POc article with two forms, *\*a* and *\*na*, which was used with common non-human nouns. The function of this article was apparently to mark the noun phrase as definite. There was also an indefinite common non-human article *\*ta* (see data in Ross 1988: 357–360, as well as Mwotlap /te-/ ‘partitive’ and Anejoñ *tah* ‘indefinite singular’ in this volume). If the noun was used in a generic sense or as an attribute, no article occurred.

The distinction between the article forms *\*a* and *\*na* is intriguing. If they were allomorphs, we would expect them to survive together in widely distributed modern languages, but (apart from an artificial creation in written Fijian) we find allomorphy only two places: in quite closely related Label and the north Bougainville linkage (see Taiof, this volume), and in Ambae (north Vanuatu). The former languages share conditions of allomorphy, namely that *a* follows a consonant, *na* a vowel. However, the same rule applies to the personal article *e* and to the north Bougainville common II article *u* (see Taiof, §3.2.3), giving alternations *e~ne* and *u~nu* matching *a~na*. Although this allomorphy *may* have occurred in POc, the evidence for it is weak: it is reflected only in a few quite closely related languages and the domain of the rule is wider than could be reconstructed for POc, which lacked the article *u*. In Ambae the form *na* occurs in an object noun phrase and *a* elsewhere (the opposite of the situation reconstructed in §2.2 for pre-POc!).

If POc *\*a* and *\*na* were contrasting morphemes, we would expect to find reflexes of the contrast in a wider range of languages, but we don’t (Crowley 1985). Instead we find that some languages reflect *\*a*, others *\*na*, either as common article or as a fossilised noun-initial element. Reflexes of both occur in the Meso-Melanesian, North New Guinea, Southeast Solomonic, Central Vanuatu, and Central Pacific groups (Oceanic subgrouping is described in Ch. 5). Thus reflexes of the two forms often occur in quite closely related languages, but only one or the other, and not both, survives in any given language.

The most appropriate kind of explanation that can be offered for such a distribution is that *\*a* and *\*na* were once contrasting morphemes, but the basis for the contrast had been eliminated by changes which had occurred at a stage immediately prior to the break-up of POc, and languages have got rid of the resulting redundancy by eliminating one or the other form. Just such an explanation was offered in §2.2 by proposing that the POc common non-human articles are descended from the three PMP article-like morphemes *\*a* ‘nominative (definite)’, *\*na* ‘genitive (definite)’, and *\*ta* ‘accusative (indefinite)’. Each language dealt with the redundancy of *\*a* (intransitive subject/transitive object) and *\*na* (transitive subject) by dropping one or the other (or both). Of the two, *\*na* survived more widely. As well as the groups mentioned above, it is also found in Yapese and was the sole survivor in the Admiralties, Utupua and Vanikoro, northern Vanuatu, southern Vanuatu and New Caledonia. Somewhat embarrassing for this explanation is the fact that in Ambae *na* marks the transitive **object**, and *a* the subject.

POc demonstratives almost certainly came at the end of their noun phrase and made a three-way distinction (Ch. 3, §2.3). Syntactically, two kinds of demonstrative are widespread in modern languages: enclitics and locative pro-forms. It would be tempting to assume that the enclitics were simply locative pro-forms that had become cliticised, were it not for the fact that their forms differ from the locative proforms, and we can reconstruct *\*=ne* 'close', *\*=ta* 'at an intermediate distance', and *\*=wa* 'distant'.

Locative pro-forms could be used independently or as demonstratives. It is doubtful whether all the forms reconstructed below occurred in POc, but all are quite widely reflected in both Oceanic and non-Oceanic languages. It is probable that the alternation between forms consisting of a vowel and *\*n*-initial forms had the same origin as the article forms *\*a* and *\*na* discussed in §2.2 and above.

|                                          |               |                 |             |
|------------------------------------------|---------------|-----------------|-------------|
| 'here', 'this', 'near speaker'           | <i>*i, *e</i> | <i>*ni, *ne</i> |             |
| 'there', 'that', 'near hearer'           | <i>*a</i>     | <i>*na</i>      | <i>*ri</i>  |
| 'yonder', 'away from both interlocutors' | <i>*o, *u</i> | <i>*no, *nu</i> | <i>*rai</i> |

The locative pro-forms *\*a* and *\*na* were apparently formally identical with the articles *\*a* and *\*na*. This may reflect the grammaticisation of the pro-forms as articles at some stage in the far past. The grammaticisation of locatives or demonstratives as articles seems reasonably common in the world's languages, the contrast being maintained by stress on locatives and demonstratives but not on articles (cf. Himmelmann 1997).

### 3.2.4 Numerals and number-marking

The POc numeral morphemes were:

|                                      |                          |
|--------------------------------------|--------------------------|
| 1 <i>*ta-sa, *sa-kai, *tai, *kai</i> | 6 <i>*onom</i>           |
| 2 <i>*rua</i>                        | 7 <i>*pitu</i>           |
| 3 <i>*tolu</i>                       | 8 <i>*walu</i>           |
| 4 <i>*pati, *pat</i>                 | 9 <i>*siwa</i>           |
| 5 <i>*lima</i>                       | 10 <i>*sa[-ŋa]-puluq</i> |

The interrogative numeral 'how many?' was *\*pica(n)*. The decades above 10 were *\*rua-ŋa-puluq* '20', *\*tolu-ŋa-puluq* '30' etc. Numeral forms between the decades were constructed with a conjunction, probably *\*ma* 'and', e.g. *\*rua-ŋa-puluq ma tolu* '23'. 'A hundred' was *\*Ratu(s)*. However, to judge from today's traditional cultures, numerals above ten were not much used in parts of early Oceania. Indeed, so widespread are quinary systems (but with the term for 10 preserved) that one suspects that the numerals 6–9 were dropping out of use among some early Oceanic speakers. Tolai (New Ireland) and Merei (north Vanuatu) reflect the quinary numerals *\*(l,r)apo-rua* '7', *\*(l,r)apo-tolu* '8', *\*(l,r)apo-pat[i]* '9'.<sup>12</sup>

POc numerals from 2 upwards seem to have had a double status: they were both adjectival verbs (§3.2.5) and nouns. They retain adjectival verb status or a fossilised reflex of it in the shape of a 3SG subject prefix in scattered Oceanic languages (e.g. Bali-Vitu, Siar, Gela, Paamese, Marquesan and Iaai). We find reflexes of the numeral as a main verb in a standard clause (§2.2):

|                            |          |              |
|----------------------------|----------|--------------|
| <i>*i=tolu</i>             | <i>a</i> | <i>Rumaq</i> |
| 3SG=three                  | ART      | house        |
| '(there are) three houses' |          |              |

We also find reflexes of it as an unmarked relative clause, as in *\*na Rumaq i-tolu*, and, as we would expect of an adjectival verb, as an unaffixed postmodifier, as in *\*na Rumaq tolu*. The nominal status of numerals is seen in the construction below (with reflexes in Yapese, New Ireland languages, Kwaio and Mwotlap), where *\*tolu* is head of the noun phrase, followed by the non-specific possessive particle *\*qi* (§3.2.7.1, Ross 1998c):

|                                            |           |              |
|--------------------------------------------|-----------|--------------|
| <i>tolu</i>                                | <i>qi</i> | <i>Rumaq</i> |
| three                                      | <i>qi</i> | house        |
| 'three houses' (= 'a threesome of houses') |           |              |

The reconstruction of the POc term(s) for 'one' is more difficult, as several morphemes and morpheme combinations occur. These morphemes presumably differed semantically and functionally in much the same way as English *a/an*, *one*, *single*, and *only*, but these differences have yet to be sorted out. The morpheme *\*sa*, occurring in 10, reflects PMP *\*esa* 'one' and is reflected in the Proto Polynesian indefinite singular article *\*ha*, but was apparently often bound in POc, where *\*ta-sa* and *\*sa-kai* are also reconstructible. We take *\*ta* in *\*ta-sa* to be the indefinite article (§3.2.3); reflexes of *\*ta* are used alone as numeral 'one' in a number of WOc languages, but we take this to be a derived usage. There are also widespread reflexes of *\*tai* 'one' but in the absence of evidence to the contrary we take this to be a morpheme separate from *\*ta*; we also take *\*tewa* 'one', reflected in languages of north and central Vanuatu and the Central Pacific to include a reflex of *\*tai*. Reflexes of the morpheme *\*kai* of *\*sa-kai* also occur alone and in *\*ti-kai* as the numeral 'one', and *\*ti* in its turn is reflected alone for 'one' in the Huon Gulf family.

A scattering of Oceanic languages (as well as non-Oceanic languages in Indonesia) use a classifier with a numeral, whilst others have fossilised reflexes of classifiers. Classifier languages include the Admiralties family, the Kilivila family, Sudest (Papuan Tip linkage), the North Bougainville linkage, the Cristobal–Malaitan languages (Southeast Solomonian), the Nuclear Micronesian family, languages in New Caledonia, and Polynesian. Generally, the numeral and classifier are bound to each other in one or the other order, and this has given rise to fossils like Tigak (New Ireland) *potul*, Toga (North Vanuatu) *vutal* both 'three'. However, it is probable that POc classifiers were not bound forms, but nouns (as in Indonesian languages), and that the different bound orders have arisen from the different structures illustrated above (see also Ross 1998c). In Sudest, nouns may still be used as counting classifiers. There is ample evidence that the general classifier was *\*puaq*, literally 'fruit', and so phrases containing a numeral like the following examples from Tigak and Mwotlap

|                          |                                   |
|--------------------------|-----------------------------------|
| Tigak:                   | Mwotlap:                          |
| <i>ta potul a nik</i>    | <i>na-tam<sup>w</sup>an vo-yo</i> |
| ART three LINKER coconut | ART-man vo-two                    |
| 'three coconuts'         | 'two men'                         |

were derived from a POc structure with a classifier like the following:

|                             |                                        |
|-----------------------------|----------------------------------------|
| <i>*ta puaq tolu a niuR</i> | <i>*na tam<sup>w</sup>ane puaq rua</i> |
| ART fruit three ART coconut | ART man fruit two                      |
| 'three coconuts'            | 'two men'                              |

In each of these POc reconstructions, a classifier + numeral phrase is apparently in apposition to the entity being counted. In all probability classifiers were not obligatory in POc, but common, as a word like *\*niuR* 'coconut' depended on a classifier for its

disambiguation, its meaning embracing the tree, its fruit, and the contents of the fruit. Other reconstructible classifiers are *\*m<sup>w</sup>ane* 'animate being' (from *\*m<sup>w</sup>aqane* 'man'), *\*kaiu* 'wooden or elongated object' (= 'tree'), *\*tau* 'person'.

There is also a classifier construction with the non-specific possessive particle, but clear reflexes are limited to CEOc languages (Hooper 1985, Ross 1998c):

*\*i-tolu puaq qi pudi*  
 3SG-three fruit *qi* banana  
 'three bananas' (= '(there are) three fruit of banana')

When someone counted objects 'one, two, three...', the numerals were prefixed with *\*ka-*, e.g. *\*ka-rua*, *\*ka-tolu*. The counting form for 'one' was apparently *\*kesa* (from PMP *\*ka-esa*). Reflexes of these forms are used for ordinals in some Oceanic languages, but there is good evidence (from Nakanai, Bugotu, Kwaio and Tamabo) that POc ordinals were formed with *\*-ña* and used as the head of a noun phrase in an appositional structure (this *\*-ña* may be the 3SG possessive suffix, but the data do not allow us to be sure):

*\*a tolu-ña a boRok*  
 ART three-ORDINAL ART pig  
 'the third pig' (cf. *\*a tolu-ña* 'the third one')

Multiplicatives were formed with the causative prefix *\*pa[ka]-* (§3.3.1) and were evidently verbs which could be used alone, e.g. *\*i-pa[ka]-tolu* (3SG-CAUS-three) 's/he did it three times', or as the second verb in an ambient serial construction (Ch. 3, §3.3.3). Distributives were formed by reduplication, e.g. *\*i-tolu-tolu* (3SG-three-three) 'three by three', and also served as the second verb in an ambient serial construction.

As we noted in §3.2.2, plurals of some common human nouns were probably formed optionally by reduplication. Where it was otherwise necessary to mark a noun as plural, one of two strategies seems to have been used. For common human nouns, a non-singular independent pronoun (§3.2.1) was preposed: *\*[k]ira tinoni* (3PL person) 'the people'. For common nouns, there was a slot immediately after the article which, to judge from modern languages, could be occupied by various morphemes (only some marking number), of which three are currently reconstructible: *\*palu* 'some, several', *\*mana* 'plural', and *\*tiqi* 'diminutive'.<sup>13</sup> Hence we can reconstruct, e.g., *\*a palu boRok* (ART some pig) 'several pigs'.

### 3.2.5 Adjectives and nominal modifiers

POc had two classes of forms which functioned as attributive modifiers of nouns: adjectival verbs and adjectival nouns (Ross 1998a). The adjectival verb class was large and to all intents and purposes an open class. Its members (which included numerals; §3.2.4) functioned both as stative verbs and as postnominal modifiers, e.g.,

|                        |                     |
|------------------------|---------------------|
| <i>*i-mataq a wai</i>  | <i>*a wai mataq</i> |
| 3SG-unripe ART mango   | ART mango unripe    |
| 'The mango is unripe.' | 'the unripe mango'  |

Adjectival nouns comprised a small, closed class of attributive postmodifiers which roughly matched the archetypal small adjective category of Dixon (1977). They included lexemes of dimension, age and value, and functioned both as heads of noun phrases and as postnominal modifiers. Whereas an adjectival verb behaved as a stative

verb when it was used predicatively, an adjectival noun was head of a predicate noun phrase:

|                               |                        |
|-------------------------------|------------------------|
| <i>*a paqoRu a Rumaq</i>      | <i>*a Rumaq paqoRu</i> |
| ART new.one ART house         | ART house new.one      |
| 'The house is (a) new (one).' | 'the new house'        |

Reflexes of adjectival nouns suggest that they underwent two forms of reduplication. Reduplication of the first one or two syllables expressed intensity. Reduplication of just the stressed syllable expressed the plural, a pattern also found with human nouns (§3.2.2).

The POc suffix *\*-[k]a* seems to have derived adjectival nouns from other roots, usually nouns (Ross 1997b), *\*-ka* occurring on vowel-final roots, *\*-a* on consonant-final. Its reflexes remain productive in modern languages, and are particularly common in colour terms, e.g., in Tamabo (this volume):

|             |                |                |          |
|-------------|----------------|----------------|----------|
| <i>dae</i>  | 'blood'        | <i>dae-ya</i>  | 'red'    |
| <i>jori</i> | 'yellow fever' | <i>jori-ya</i> | 'yellow' |

Many WOc languages have adjectival classes and syntax quite different from those described here, but these represent innovations which occurred largely after the break-up of POc (Ross 1998b).

### 3.2.6 Basic noun phrase structure

All the elements of the basic structure are discussed above:

| ART +      | (PREMODIFIER +)          | NOUN (+ | MODIFIER) (+    | DEMONSTRATIVE) |
|------------|--------------------------|---------|-----------------|----------------|
| §3.2.3     | §3.2.4                   | §3.2.2  | §3.2.5          | §3.2.3         |
| <i>*a</i>  | e.g. <i>*palu</i> 'some' |         | adjectival verb |                |
| <i>*na</i> |                          |         | adjectival noun |                |
| <i>*ta</i> |                          |         |                 |                |

Two or more noun phrases were linked in a co-ordinate noun phrase by a morpheme functioning as a conjunction. This was often *\*ma* 'and', which seems normally to have had the transitive suffix *\*-i* (hence its reflex is *me* in many languages) and an object enclitic agreeing in number with the following noun phrase, i.e. *\*X ma-i=a Y* (where *Y* is singular). The morphology indicates that it was (originally, at least) a preposition rather than a conjunction.

### 3.2.7 Possession

POc had direct and indirect possessive constructions (Ch. 3, §2.7). Often, as in modern languages, the possessor was indicated by the possessor suffix (§3.2.1) alone. Where the possessor was named, the possessor noun phrase followed the rest of the construction (Lichtenberk 1985). If the possessor noun was a proper name, then it was evidently introduced by the personal article *\*i* (§3.2.3), which was cliticised to the final vowel of the preceding element (see examples below and Hooper 1985). If the possessor was a common human noun, it had no article. If it was common non-human, we assume that the article was *\*na* (from the PMP genitive article), although there is no formal evidence for *\*na* rather than *\*a* (§3.2.3).

## 3.2.7.1 Direct possession

In the direct construction, the person and number of the possessor were marked by a possessor suffix: POc *\*a qaqe-ña* 'her/his leg', *\*a qaqe-gu* 'my leg', *\*a qaqe-mu* 'your (singular) leg', etc. With a possessor noun (rather than a pronoun) this became:

|                            |                    |
|----------------------------|--------------------|
| <i>*a qaqe-ña na boRok</i> | <i>*a qaqe=i X</i> |
| ART leg-3SG ART pig        | ART leg=ART X      |
| 'the pig's leg'            | 'X's leg'          |

As noted in Ch. 3, the direct/indirect structural distinction generally reflects the inalienable/alienable semantic distinction. Directly possessed nouns in POc probably included most body parts, most kin terms, and most locative parts, but we will suggest in §3.2.7.2 that there was considerable potential for a noun to move back and forth between the direct and indirect constructions. It also seems very likely that the semantic category of passive possession was expressed by the direct possession construction in POc, whereas it is assigned to the food category of indirect possession in many modern languages (Lynch 1996). Direct possession included 'possession' of (a) an inalienable property and (b) a nominalised verb by its object or beneficiary. For example:

|                                        |                               |
|----------------------------------------|-------------------------------|
| <i>*a lab<sup>w</sup>a-ña na boRok</i> | <i>*a tanum-a-dra na pudi</i> |
| ART big-3SG ART pig                    | ART bury-NOM-3PL ART banana   |
| 'the pig's size'                       | 'the planting of the bananas' |

Both these reconstructions make the assumption that a final consonant was lost before a possessor suffix: *\*lab<sup>w</sup>at* became *\*lab<sup>w</sup>a-*, *\*tanum-an* became *\*tanum-a-*. Although evidence from much of Oceania supports this assumption, data from Tanna languages (Southern Vanuatu) speak against it. Here, the few nouns with (Proto Tanna) final consonants other than *\*-n* reflect the final consonant before a possessor suffix. Thus POc *\*na icuŋ + -ña* 'ART nose 3SG' has become Lenakel *nhaŋən*, suggesting vowel insertion before suffixation (*\*na icuŋu-ña*), whereas Kairiru *isu-ñ*, Bali *iru-na* Roviana *isu-na*, Lengo *iðu-na* all reflect POc *\*icu-ña*. The best interpretation of these data is that POc allowed collocations like *\*icuŋ-ña*. Since possessor suffixes could also occur on alienable nouns like *\*Rumaq* 'house' in POc (§3.2.7.2) this pattern would have been quite common. As over time possessor suffixes came to be limited to inalienable nouns in most Oceanic languages, they also became more closely phonologically bonded to them, and the non-canonic consonant sequence in, e.g., *\*icuŋ-ña* was usually eliminated by deleting the root-final consonant, but in languages like Proto Tanna by inserting an epenthetic vowel.

If the possessor was non-specific, then the possessor suffix disappeared and was replaced by the non-specific possessive particle *\*qi* (Hooper 1985, Ross 1998c):

|                         |
|-------------------------|
| <i>*a qaqe qi boRok</i> |
| ART leg <i>qi</i> pig   |
| 'a pig's leg'           |

To judge from the scattering of languages where the adjectival nouns 'big' and 'small' respectively reflect *\*tina-* 'mother' and *\*natu-* 'child', it seems that this construction was also used in two idioms reflecting a pattern widespread in Asia (Matisoff 1991):

|                                        |                             |
|----------------------------------------|-----------------------------|
| <i>*a tina qi Rumaq</i>                | <i>*a natu qi Rumaq</i>     |
| ART mother <i>qi</i> house             | ART child <i>qi</i> house   |
| 'a whopper of a house', 'a huge house' | 'a tiny house'              |
| (lit. 'a mother of a house')           | (lit. 'a child of a house') |

## 3.2.7.2 Indirect Possession

In indirect possession, the possessor suffix was attached not to the possessed noun but to a possessive ‘classifier’, giving a structure like:

\*a na-ña Rumaq  
ART CL-3SG house  
‘her/his house’

\*a na-gu Rumaq  
ART CL-1SG house  
‘my house’

The fact that the classifier–suffix sequence almost certainly preceded the possessed noun (Lichtenberk 1985) implies that it was the head of the construction and that the possessed noun was (originally, at least) in apposition to it.

With a possessor noun phrase this apparently became:

\*a Rumaq na-ña na boRok  
ART house CL-3SG ART pig  
‘the pig’s house’

\*a Rumaq na-ña tam<sup>w</sup>ata  
ART house CL-3SG man  
‘the man’s house’

\*a Rumaq na=i X  
ART house CL=ART X  
‘X’s house’

Note that this entailed an exchange in the relative positions of the classifier–suffix sequence and the possessed noun. The alternate structure below probably also occurred, the difference between the two being one of information structure:

\*a na-ña Rumaq tam<sup>w</sup>ata  
ART CL-3SG house man  
‘the man’s house’

We have no evidence as to the morphology of the corresponding structure with a proper name.

If the possessor was non-specific, then the classifier–suffix sequence was replaced by the morpheme \*ni (descendent of PMP \*ni indefinite genitive article; §2.2). In cases like this the ‘possessor’ was semantically often an attribute of one kind or another:

\*a polo ni niuR  
ART juice ni coconut  
‘coconut water’

\*a Rumaq ni turuR  
ART house ni sleep  
‘a rest house’ (lit. ‘house of sleep’)

\*a Rumaq ni Reqi  
ART house ni elephant.grass  
‘a grass house’

The indirect possessive constructions above all contain the reconstructed classifier \*na. There is broad agreement among scholars that three POC classifiers can be reconstructed: \*ka- ‘food’, \*m(ʷ)a- ‘drink’ (perhaps in fact \*m<sup>w</sup>a-) and \*na- ‘general’, the last-named including any indirect possession relationship not included under food or drink. Whilst these reconstructions are almost certainly justified, with almost equal certainty they do not represent an exhaustive listing of POC possessive classifiers. First, there are reflexes not only of \*na- but also of \*a-, an alternation reminiscent of the