

The lexicon of Proto Oceanic

**The culture and environment of ancestral
Oceanic society**

3 Plants

edited by

Malcolm Ross, Andrew Pawley and Meredith Osmond



Pacific Linguistics

Research School of Pacific and Asian Studies
The Australian National University

Published by Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
Canberra ACT 0200
Australia

Copyright in this edition is vested with Pacific Linguistics

First published 2008

National Library of Australia Cataloguing-in-Publication entry:

- Title: The lexicon of Proto Oceanic. The culture and environment of ancestral Oceanic society / Volume 3, Plants / editors, Malcolm Ross, Andrew Pawley and Meredith Osmond.
- ISBN: 9780858835894 (v. 3 : pbk.)
- Notes: Bibliography.
- Subjects: Proto-Oceanic language.
Ethnology—Oceania.
Oceania—Social life and customs.
- Other Authors/
Contributors: Ross, Malcolm (Malcolm D.)
Pawley, Andrew
Osmond, Meredith
Evans, Bethwyn.
Australian National University. Research School of Pacific and Asian Studies. Pacific Linguistics.
- Dewey Number: 499.4

Copyedited by Meredith Osmond and Bethwyn Evans
Maps and diagrams by Malcolm Ross
Typeset by Malcolm Ross using XeTeX and the Memoir class of LaTeX
Cover design by Julie Manley/Addcolour Digital Pty Ltd
Cover photograph by R. Michael Bourke
Printed and bound by Addcolour Digital Pty Ltd, Fyshwick, Canberra

Contents overview

Chapter contents in detail		vii
List of Tables		xv
List of Figures		xvi
List of maps		xix
Abbreviations		xxi
Acknowledgments		xxiii
Chapter 1	Introduction Malcolm Ross, Andrew Pawley and Meredith Osmond	1
Chapter 2	Introducing Proto Oceanic plant names Malcolm Ross	25
Chapter 3	Ethnobotanical classification Bethwyn Evans	53
Chapter 4	Parts of plants Bethwyn Evans	85
Chapter 5	Wild plants of the coastal strand Malcolm Ross	129
Chapter 6	Wild plants of the mangrove swamp Malcolm Ross	173
Chapter 7	Wild plants of primary lowland tropical rain forest Malcolm Ross	185
Chapter 8	Wild plants of secondary lowland rain forest and grassland Malcolm Ross	237
Chapter 9	Staple foods: root crops, bananas, breadfruit and sago Malcolm Ross	255

Chapter 10	Green vegetables and figs Malcolm Ross	293
Chapter 11	Nut and fruit trees Malcolm Ross	311
Chapter 12	The coconut palm Malcolm Ross and Bethwyn Evans	355
Chapter 13	Other cultivated plants Malcolm Ross	389
Chapter 14	Concluding notes Malcolm Ross	427
Appendix A	Data sources and collation	437
Appendix B	Languages	443
References		471
Index of reconstructions by protolanguage		493
Alphabetical index of reconstructions		507
Index of plant names		523
Index of plant names by families		533

Chapter contents in detail

Contents overview	v
List of Tables	xvi
List of Figures	xvii
List of maps	xx
Abbreviations	xxi
Acknowledgments	xxiii
1 Introduction	1
1 Aims	1
2 The relation of the current project to previous work	2
3 Reconstructing the lexicon	4
3.1 Terminological reconstruction	4
3.2 Subgrouping and reconstruction	6
3.2.1 Subgrouping	6
3.2.2 Kinds of subgroup	7
3.2.3 Criteria for reconstruction	10
3.2.4 Further notes on subgroups	13
3.3 Sound correspondences	15
3.4 POc phonology and orthography	16
4 Conventions common to the series	17
4.1 Chapter format	17
4.2 Data	19
4.3 Conventions used in representing reconstructions	20
5 Data sources and conventions specific to this volume	22
5.1 Cognate sets and reconstructions	22
5.2 Botanical and ethnographic information and vernacular names	22
5.3 Scientific plant names	23
5.4 English and Pacific pidgin plant names	24
5.5 Indexes	24

2	Introducing Proto Oceanic plant names	25
1	Floristic regions	25
2	Ways of classifying plants	25
2.1	Classification by use	26
2.2	Classification by ethnobotanical category	26
2.3	Classification by vegetation habitat and cultivation status	27
3	Vegetation habitats in NW Island Melanesia	28
3.1	Natural vegetation habitats	28
3.1.1	Coastal strand vegetation	28
3.1.1.1	Herbaceous zone	28
3.1.1.2	Beach scrub	29
3.1.1.3	Littoral forest	29
3.1.2	Mangrove forest	29
3.1.3	Freshwater swamp forest	30
3.1.4	Primary lowland rain forest	30
3.2	Anthropogenic vegetation habitats	30
3.2.1	Garden vegetation	30
3.2.2	Secondary lowland rain forest	31
3.2.3	Grassland	32
4	The relationship between garden and bush in Proto Oceanic society	32
5	Proto Oceanic terms for types of vegetation and land use	34
6	The ingredients of an Oceanic meal	36
6.1	Starchy food	40
6.2	Additional ingredients	43
7	Apparent encodings of plant categorisations in Oceanic plant names	43
7.1	Prefixes	44
7.1.1	Reflexes of POc <i>*kayu</i> ‘tree, shrub’	44
7.1.2	POc <i>*ka-</i> in trisyllabic tree names	46
7.1.3	Reflexes of POc <i>*puqu(n)</i> ‘tree, shrub’ and <i>*raun</i> ‘leaf’	47
7.1.4	POc <i>*mala-</i> ‘resembling’	49
7.2	Reduplication	50
3	Ethnobotanical classification	53
1	Introduction	53
2	Ethnobiological classifications	53
3	Botanical taxonomies in modern Oceanic languages	58
3.1	Wayan Fijian	58
3.2	Kwara’ae	63
3.3	Arosi	67
3.4	Nduke	69
3.5	Samoan	69
4	Botanical taxa in Proto Oceanic	70
4.1	Plant	70
4.2	Tree	71
4.3	Palms	73
4.4	Vine, creeper	74

4.5	Grass	75
4.6	Moss, algae, seaweed	76
4.7	Mushrooms, fungi	78
4.8	Other terms	79
5	Conclusions	82
4	Parts of plants	85
1	Plant part terms in modern Oceanic languages	85
2	Plant part terms reconstructable for POc	89
2.1	Stems of woody plants	89
2.2	Stems of non-woody plants	92
2.3	Branches	93
2.4	Roots	97
2.5	Leaves	103
2.6	Shoots, sprouts and suckers	108
2.7	Flowers	112
2.8	Fruit	115
2.9	Seeds	117
2.10	Outer coverings	120
2.11	Sap, resin	123
2.12	Thorns	124
3	Concluding remarks	126
5	Wild plants of the coastal strand	129
1	Introduction	129
2	Coastal shallows	130
3	The herbaceous zone	131
3.1	<i>Ipomoea</i> spp., morning glory	131
3.2	<i>Wedelia biflora</i>	133
4	Beach scrub	134
4.1	Woody shrubs and small trees	134
4.1.1	<i>Cordia subcordata</i> , kerosene wood, island walnut, sea trumpet, Pacific rosewood	134
4.1.2	<i>Hernandia nymphaefolia</i> , sea hearse tree, lantern tree	136
4.1.3	<i>Hibiscus tiliaceus</i> , beach hibiscus	138
4.1.4	<i>Pemphis acidula</i>	140
4.1.5	<i>Scaevola taccada</i>	141
4.1.6	<i>Thespesia populnea</i>	143
4.1.7	<i>Tournefortia argentea</i> , tree heliotrope, beach heliotrope	144
4.1.8	<i>Vitex trifolia</i>	145
4.2	Climbers	147
4.2.1	<i>Flagellaria indica</i> , supplejack	147
4.2.2	<i>Hoya</i> spp.	148
5	Littoral forest	148
5.1	<i>Adenanthera pavonina</i> , bead tree, red sandalwood	148
5.2	<i>Barringtonia asiatica</i> , sea poison tree, fish poison tree	149

5.3	<i>Calophyllum inophyllum</i> , portia, Indian laurel, Alexandrian laurel, beach mahogany	152
5.4	<i>Casuarina equisetifolia</i> , casuarina, beach she-oak	156
5.5	<i>Erythrina variegata</i> , coral tree, Indian coral tree	158
5.6	<i>Fagraea berteroana</i>	161
5.7	<i>Flueggea flexuosa</i>	164
5.8	<i>Guettarda speciosa</i>	165
5.9	<i>Gyrocarpus americanus</i> , canoe tree	166
5.10	<i>Neisosperma oppositifolium</i> , twin apple	166
5.11	<i>Pisonia</i> spp.	167
5.12	<i>Pongamia pinnata</i>	170
5.13	<i>Premna</i> spp.	171
6	Wild plants of the mangrove swamp	173
1	Introduction	173
2	The <i>Rhizophora</i> zone of the mangrove forest	174
2.1	<i>Rhizophora</i> spp.	174
3	The <i>Bruguiera</i> zone of the mangrove forest	176
3.1	<i>Bruguiera</i> spp.	176
4	The landward margin of the mangrove forest	178
4.1	<i>Cerbera</i> spp.	178
4.2	<i>Dolichandrone spathacea</i>	179
4.3	<i>Excoecaria agallocha</i> , milky mangrove	180
4.4	<i>Heritiera littoralis</i>	181
4.5	<i>Nypa fruticans</i> , nipa(h) palm	182
4.6	<i>Xylocarpus</i> spp., puzzlenut tree	183
7	Wild plants of primary lowland tropical rain forest	185
1	Introduction	185
2	Rain forest layers	186
3	Emergent trees	187
3.1	<i>Alstonia scholaris</i> , milky pine, canoe tree	187
3.2	<i>Falcataria moluccana</i>	188
3.3	<i>Octomeles sumatrana</i>	189
4	The forest canopy	190
4.1	<i>Camptosperma brevipetiolatum</i>	190
4.2	<i>Cinnamomum</i> spp., wild cinnamon	191
4.3	<i>Dillenia schlechteri</i>	193
4.4	<i>Dracontomelon dao</i> , New Guinea walnut, Dragon plum	194
4.5	<i>Dysoxylum</i> spp., stinktree, stinkwood	195
4.6	<i>Elaeocarpus</i> spp.	197
4.7	<i>Endospermum</i> spp., cheesewood, milkwood, whitewood, PNG basswood	197
4.8	<i>Garuga floribunda</i>	199
4.9	<i>Intsia bijuga</i> , ironwood, island teak	200
4.10	<i>Planchonella</i> spp.	203
4.11	<i>Pterocarpus indicus</i> , New Guinea rosewood, flame wood	204

4.12	<i>Vitex cofassus</i> , New Guinea teak	205
5	The lower canopy	207
5.1	<i>Bischofia javanica</i> , Java cedar	207
5.2	<i>Cananga odorata</i> , ylang-ylang, perfume tree	209
5.3	<i>Cryptocarya</i> sp.	209
5.4	<i>Dillenia</i> sp.	210
5.5	<i>Diospyros</i> spp., ebony, blackwood	211
5.6	<i>Euodia</i> spp.	211
5.7	<i>Kleinhovia hospita</i> , puzzle tree, guest tree	212
5.8	<i>Litsea</i> spp.	215
5.9	<i>Myristica</i> spp., wild nutmeg	215
5.10	<i>Parinari</i> spp., putty nut	217
5.11	Palms	221
5.11.1	<i>Caryota rumphiana</i> , black palm, fishtail palm	221
5.11.2	<i>Licuala</i> spp., fan palm	222
6	The shrub layer	223
6.1	Woody shrubs	223
6.1.1	<i>Abroma augusta</i>	223
6.1.2	<i>Angiopteris evecta</i> , mule's foot fern	224
6.1.3	<i>Donax cannaeformis</i>	224
6.1.4	<i>Garcinia</i> spp.	225
6.1.5	<i>Phaleria</i> spp.	226
6.1.6	<i>Semecarpus forstenii</i> , poisonwood	227
6.2	Leafy shrubs	228
6.2.1	<i>Hornstedtia lycostoma</i>	228
6.3	Climbers and epiphytes	229
6.3.1	<i>Calamus</i> spp., rattan, lawyer cane	229
6.3.2	<i>Dendrocnide</i> and <i>Laportea</i> spp., nettle trees	230
6.3.3	<i>Lygodium</i> spp.	233
6.3.4	<i>Merremia</i> spp.	234
6.3.5	<i>Asplenium nidus</i> , bird's nest fern	234
7	The forest floor	235
8	Wild plants of secondary lowland rain forest and grassland	237
1	Introduction	237
2	Trees and shrubs	238
2.1	<i>Acalypha</i> spp., copper leaf	238
2.2	<i>Alphitonia</i> spp.	239
2.3	<i>Commersonia bartramia</i>	240
2.4	<i>Glochidion philippicum</i> , little cheese tree	241
2.5	<i>Macaranga</i> spp.	242
2.6	<i>Pipturus argenteus</i>	244
2.7	<i>Rhus taitensis</i>	246
2.8	<i>Trema orientalis</i>	247
2.9	<i>Trichospermum</i> spp.	248
3	Grasses	249

3.1	<i>Imperata cylindrica</i> , sword grass, blady grass, cogon grass	250
3.2	<i>Coix lachryma-jobi</i> , Job's tears	252
3.3	<i>Miscanthus floridulus</i>	253
3.4	<i>Saccharum spontaeum</i> , wild sugarcane	253
9	Staple foods: root crops, bananas, breadfruit and sago	255
1	Introduction	255
2	Root crops	255
2.1	Yams, <i>Dioscorea</i> spp.	256
2.1.1	Why so few yam terms?	257
2.1.2	<i>Dioscorea alata</i> , greater yam, water yam, winged yam	258
2.1.3	<i>Dioscorea bulbifera</i> , potato yam, aerial yam, air potato	262
2.1.4	<i>Dioscorea esculenta</i> , lesser yam, prickly yam	263
2.1.5	Other terms for yam	264
2.2	Taro (Araceae)	265
2.2.1	<i>Colocasia esculenta</i> , taro, Asian taro	265
2.2.2	<i>Cyrtosperma merkusii</i> , swamp taro, giant taro	269
2.2.3	<i>Alocasia macrorrhizos</i> , giant taro, elephant ear taro	271
2.2.4	Another term for taro?	273
2.3	Other tubers	273
2.3.1	<i>Amorphophallus paeoniifolius</i> , wild arrowroot, elephant yam	273
2.3.2	<i>Tacca leontopetaloides</i> , Polynesian arrowroot	273
2.3.3	<i>Pueraria lobata</i>	274
3	Bananas, cultivars of the genus <i>Musa</i>	275
4	<i>Artocarpus altilis</i> , breadfruit	280
5	Plants which store starch in their stems	286
5.1	<i>Metroxylon</i> spp., sago palm	286
5.2	<i>Cycas</i> spp., cycad, Malayan palm fern	289
6	Summary	292
10	Green vegetables and figs	293
1	Introduction	293
2	Leafy vegetables	294
2.1	<i>Abelmoschus manihot</i> , slippery cabbage, bush cabbage, island cabbage	294
2.2	<i>Amaranthus tricolor</i> , Chinese spinach, Joseph's coat amaranth	296
2.3	<i>Gnetum gnemon</i>	297
2.4	<i>Polyscias</i> spp., panax	298
3	Other green vegetables	300
3.1	<i>Cyathea</i> spp.	300
3.2	<i>Saccharum edule</i> , 'coastal pitpit' (in Papua New Guinea)	301
4	<i>Ficus</i> spp., banyans, figs (Moraceae)	302
4.1	Terms for classes of <i>Ficus</i>	302
4.2	<i>Ficus copiosa</i> and <i>F. wassa</i> , sandpaper cabbage, sandpaper fig	306
4.3	<i>Ficus tinctoria</i> , dye fig	308
4.4	Other names for <i>Ficus</i> species	309

11	Nut and fruit trees	311
1	Introduction	311
2	Domesticated nut trees	312
2.1	<i>Canarium</i> spp.	312
2.2	<i>Inocarpus fagifer</i> , Tahitian chestnut, Pacific chestnut, Polynesian chestnut	317
2.3	<i>Barringtonia</i> spp. other than <i>B. asiatica</i> , cutnut, bush nut	319
2.4	<i>Terminalia</i> spp.	323
2.5	<i>Pandanus</i> spp., screw pines	327
2.5.1	<i>Pandanus tectorius</i> , coastal pandanus	327
2.5.2	<i>Pandanus dubius</i>	330
2.5.3	<i>Pandanus conoideus</i> , red fruit pandanus	331
2.5.4	<i>Pandanus lamekotensis</i>	333
2.5.5	Other names for <i>Pandanus</i> species	333
2.6	<i>Sterculia vitiensis</i>	334
2.7	<i>Pangium edule</i> , pangi, ankle-rattle tree	335
3	Domesticated fruit trees	336
3.1	<i>Burckella obovata</i> , red silkwood	336
3.2	<i>Clymenia polyandra</i> and <i>Citrus</i> spp. (Rutaceae)	338
3.3	<i>Corynocarpus cribbianus</i>	340
3.4	<i>Mangifera</i> spp., mango	341
3.5	<i>Pometia pinnata</i> , ocean lychee, island lychee	342
3.6	<i>Spondias cytherea</i> , golden apple, Polynesian plum, hog plum, Tahitian apple	344
3.7	<i>Syzygium malaccense</i> , Malay apple, rose apple	348
4	Other nut and fruit trees	350
4.1	<i>Finschia chloroxantha</i>	351
4.2	<i>Parartocarpus venenosa</i>	351
12	The coconut palm	355
1	Introduction	355
2	The coconut palm	355
2.1	<i>Cocos nucifera</i> , coconut	355
3	Growth stages of the coconut	357
3.1	Growth stage terminologies	357
3.2	Growth stage 2: very small newly formed fruit	362
3.3	Growth stage 6: green, drinkable	363
3.4	Growth stage 9: ripe, flesh hardened	363
3.5	Growth stage 12: sprouting	366
3.6	Other growth stages	367
4	Parts of the coconut fruit	368
4.1	Coconut water	369
4.2	Coconut flesh and its products	370
4.3	The embryo and the 'apple'	372
4.4	Coconut husk	376
4.5	Coconut shell	377
5	Parts of the coconut palm	378

5.1	The frond and its parts	379
5.1.1	Green fronds	379
5.1.2	Dried fronds	381
5.1.3	Midrib of the frond	383
5.1.4	Spine/stalk of leaflet	384
5.2	The frond netting	384
5.3	The spathe, the inflorescence and the spadix	386
6	Conclusion	387
13	Other cultivated plants	389
1	Introduction	389
2	Plants grown for human consumption	389
2.1	<i>Saccharum officinarum</i> , sugarcane	389
2.2	Stimulants and narcotics	391
2.2.1	<i>Areca catechu</i> , betel palm, betelnut	391
2.2.2	<i>Piper betle</i> , betel pepper	394
2.2.3	<i>Piper methysticum</i> , kava	395
3	Plants which provide materials for making things	398
3.1	<i>Bambusa</i> , <i>Schizostachyum</i> and <i>Nastus</i> spp., bamboo	398
3.2	<i>Aleurites moluccana</i> , candlenut	403
3.3	<i>Broussonetia papyrifera</i> , paper mulberry	405
3.4	<i>Morinda citrifolia</i> , Indian mulberry, yellow wood	406
4	Fish poisons	409
4.1	<i>Derris</i> spp., derris root, fish poison plant	409
5	Multi-purpose roots	411
5.1	<i>Curcuma longa</i> , turmeric	412
5.2	<i>Zingiber</i> spp., ginger	413
6	Plants with large colourful leaves	415
6.1	<i>Codiaeum variegatum</i> , croton	415
6.2	<i>Cordyline fruticosa</i> , cordyline, victory leaf, ti plant	416
6.3	<i>Dracaena angustifolia</i>	416
6.4	Reconstructions	417
6.5	<i>Heliconia indica</i> , heliconia	421
6.6	<i>Crinum asiaticum</i> , spider lily	421
7	Cucurbits	423
7.1	<i>Lagenaria siceraria</i> , bottle gourd	423
7.2	<i>Benincasa hispida</i> , wax gourd, white gourd, winter melon	423
7.3	<i>Cucurbita moschata</i> , pumpkin	424
7.4	<i>Cucumis sativus</i> , cucumber	425
14	Concluding notes	427
1	The uses of plants	427
2	The durability of Proto Oceanic plant names	428
3	Where do Oceanic plant names come from?	428
Appendix A	Data sources and collation	437
1	Data sources	437

1.1	General sources	437
1.1.1	Published sources	437
1.1.2	Unpublished sources	439
1.2	Plant name sources	440
2	Linnaean terms	441
3	Collation	442
Appendix B Languages		443
1	Introduction	443
2	Languages by subgroups	443
3	Language finderlist	455
Index of reconstructions by protolanguage		493
Alphabetical index of reconstructions		507
Index of plant names		523
Index of plant names by families		533

1 *Introduction*

MALCOLM ROSS, ANDREW PAWLEY AND MEREDITH OSMOND

1 Aims

This is the third in a series of volumes bringing together the results of recent work on the lexicon of the Proto Oceanic (POc) language.¹ POc is the immediate ancestor of the Oceanic subgroup of the Austronesian language family (see Map 1). This subgroup consists of all the Austronesian languages of Melanesia east of 136° E, together with those of Polynesia and (with two exceptions) those of Micronesia—more than 450 languages in all.² Extensive arguments for the existence of Oceanic as a clearly demarcated branch of Austronesian were first put forward by Otto Dempwolff in the 1920s, and the validity of the subgroup is now recognised by virtually all scholars working in Austronesian historical linguistics.

The development and break-up of the POc language and speech community were stages in a truly remarkable chapter in human prehistory—the colonisation by Austronesian speakers of the Indo-Pacific region in the period after about 2000 BC. The outcome was the largest of the world's well-established language families and (until the expansion of Indo-European after Columbus) the most widespread. The Austronesian family comprises more than 1,000 distinct languages. Its eastern and western outliers, Madagascar and Easter Island, are two-thirds of a world apart, and its northernmost extensions, Hawaii and Taiwan, are separated by 70 degrees of latitude from its southernmost outpost, Stewart Island in New Zealand.

It is likely that the divergence of Oceanic from its nearest relatives, which are the Austronesian languages spoken around Cenderawasih Bay and in South Halmahera (Blust 1978*a*), began when Austronesian speakers from the Cenderawasih Bay area moved eastwards along the north coast of New Guinea and into the Bismarck Archipelago. There is a strong school of opinion that associates the subsequent break-up of POc with the rapid colonisation of Island

¹ The project has been jointly directed by Andrew Pawley and Malcolm Ross, with research assistance from Meredith Osmond, in the Department of Linguistics, Research School of Pacific and Asian Studies at the Australian National University. Bethwyn Evans, who has contributed to the present volume, was employed as a Research Associate under a grant to the project from the Australian Research Council. Originally, five volumes were planned, but the large amount of material on plants and animals has led to the splitting of the planned volume 3 into two.

² The listing in Tryon (1995) contains 466 Oceanic languages, many of which are subdivisible into dialects.

Melanesia and the central Pacific by bearers of the Lapita culture between about 1200 and 900 BC (see Map 2 and volume 2, chapter 2).

The present project aims to bring together a large corpus of lexical reconstructions for POc, with supporting cognate sets, organised according to semantic fields and using a standard orthography for POc. We hope that this thesaurus will be a useful resource for culture historians, archaeologists and others interested in the prehistory of the Pacific region. The comparative lexical material should also be a rich source of data for various kinds of purely linguistic research, e.g. on semantic change and subgrouping in the more than 400 daughter languages.

Volume 1 of *The lexicon of Proto Oceanic* dealt with material culture. Volumes 2, 3 and 4 examine relevant sets of cognate terms in order to gain insights into how POc speakers viewed and exploited their environment. Volume 2 dealt with the geophysical or inanimate environment, the present volume treats plants and volume 4 animals. Volume 5, as it is planned at the time of writing, will deal with terminologies centring on people and society, including the body and human conditions and activities, and social organisation, belief systems, rituals, recreation and other elements of non-material culture. Volume 6 will treat closed categories including adjectives, pronouns, and number. It will also include a review of the main findings of the project, especially comparing the evidence from archaeology and historical linguistics concerning the culture and dispersal of Austronesian speakers into and across the Pacific Islands. Volume 6 will also provide an index to the POc and other reconstructions presented in the whole work, as well as an English-to-POc finderlist and a list of all languages cited, together with their subgroups.³

The organisation of the present volume is as follows. Chapter 2 is an introduction to the present volume, discussing plant classifications and plant names. It also provides a rationale for the organisation of this volume. Chapter 3 reconstructs major categories of POc ethnobotanical classification. Chapter 4 deals with POc terms for the parts of trees and other plants. Chapters 5–8 present reconstructed names for wild trees and plants of the bush. They are organised according to vegetation habitat. Chapter 5 treats plants of the coastal strand, chapter 6 plants of the mangrove swamp, chapter 7 primary lowland rain forest plants and chapter 8 plants that are found mainly in secondary lowland rain forest and grassland. Chapters 9–12 present reconstructed names for cultivated food plants.⁴ Chapter 9 deals with staple foods, chapter 10 with green vegetables, chapter 11 with nut- and fruit-bearing trees and plants and chapter 12 with the coconut palm and its products. Finally, chapter 13 treats trees and plants that are cultivated for purposes other than food production.

2 The relation of the current project to previous work

Reconstructions of POc phonology and lexicon began with Dempwolff's pioneering work in the 1920s and 1930s. Dempwolff's dictionary of reconstructions attributed to Proto Austronesian (PAN) (1938) — but equivalent in modern terms to Proto Malayo-Polynesian (PMP) — contains some 600 reconstructions with reflexes in Oceanic languages.

³ This Introduction incorporates much of the material in the Introductions to Volume 1 and 2. We replicate this material here in order that each volume can be used independently. There are, however, some changes in the subgrouping of Oceanic languages.

⁴ Terms associated with horticultural practices are presented in chapter 5 of volume 1.

Since the 1950s, POc and other early Oceanic interstage languages have been the subject of a considerable body of research. However, relatively few new reconstructions safely attributable to POc were added to Dempwolff's material until the 1970s. In 1969 George Grace made available as a working paper a compilation of reconstructions from various sources amounting to some 700 distinct items, attributed either to POc or to early Oceanic interstages. These materials were presented in a new orthography for POc, based largely on Biggs' (1965) orthography for an interstage he called Proto Eastern Oceanic. Updated compilations of Oceanic cognate sets were produced at the University of Hawaii in the period 1977–1983 as part of a project directed by Grace and Pawley. These compilations and the supporting data are problematic in various respects and we have made only limited use of them.

Comparative lexical studies have been carried out for several lower-order subgroups of Oceanic: for Proto Polynesian by Biggs (resulting in Walsh & Biggs 1966, Biggs et al. 1970 and subsequent versions of the POLLEX file, including Biggs & Clark (1993) and Clark & Biggs (2006));⁵ for Proto Micronesian by scholars at the University of Hawaii (Bender et al. 1983, Bender et al. 2003); for Proto North-Central Vanuatu by Clark (1996, 2008);⁶ for Proto Southern Vanuatu by Lynch (1978, 1996, 2001); for New Caledonia by Ozanne-Rivierre (1992), Haudricourt & Ozanne-Rivierre (1982) and Geraghty (1989); for Proto SE Solomonian by Levy (1980) and František Lichtenberk (1988); for Proto Central Pacific by Hockett (1976), Geraghty (1983, 1986, 1996 together with a number of unpublished papers); for Proto Eastern Oceanic by Biggs (1965), Cashmore (1969), Levy & Smith (1970), and Geraghty (1990); and for Proto Central Papuan by Pawley (1975), Lynch (1978, 1980), and Ross (1994*b*).

Robert Blust of the University of Hawaii has, in a series of papers (1970, 1980, 1983–84, 1986, 1989) published extensive, alphabetically ordered, lexical reconstructions (with supporting cognate sets) for interstages earlier than POc, especially for Proto Austronesian, Proto Malayo-Polynesian and Proto Eastern Malayo-Polynesian. He has also written several papers investigating specific semantic fields (1980, 1982, 1987, 1994). At the time of writing, Blust has a major work in progress, the Austronesian Comparative Dictionary (ACD), which will bring together all his reconstructions for Proto Austronesian and lower-order stages. This is stored in electronic form at the University of Hawaii. The version to which we refer dates from 1998.

Several papers predating our project systematically investigated particular semantic domains in the lexicon of POc, e.g. Milke (1958), French-Wright (1983), Pawley (1982, 1985), Pawley & Green (1985), František Lichtenberk (1986), R. Walter (1989), and the various papers in Pawley & Ross (1994). Ross (1988) contains a substantial number of new POc lexical reconstructions, as well as proposed modifications to the reconstructed POc sound system and the orthography.

These earlier works have provided valuable points of reference, both inside and outside the Oceanic group, and we are indebted particularly to Biggs and Clark (1993), Clark (1996*a*), Lynch (2001*c*) and Blust (ACD). However, previous Oceanic lexical studies were limited both

⁵ We have mostly referred to Biggs & Clark (1993) in our work, as Clark & Biggs (2006) became available only late in the preparation of this volume. We have referred to the latter in just a few cases.

⁶ The manuscript of Clark (2008) became available only after this volume was nearing completion, and is not otherwise referred to here.

by large gaps in the data, with a distinct bias in favour of ‘Eastern Oceanic’ languages, and by the technical problems of collating large quantities of data. Although most languages in Melanesia remain poorly described, there are now many more dictionaries and extended word lists, particularly for Papua New Guinea, than there were in the 1980s. And developments in computing hardware and software now permit much faster and more precise handling of data than was possible then. A list of sources and a summary of the Project’s collation procedures is found in Appendix 1.

As the present project proceeded, we came to realise that the form in which preliminary publications were presented—namely as essays, each discussing cognate sets for a particular semantic field at some length—would also be the best form for the presentation of our final synthesis. A discursive treatment of individual terminologies, as opposed, say, to a dictionary-type listing of reconstructions with supporting cognate sets, makes it easier to relate the linguistic comparisons to relevant issues of culture history, language change, and methodology. Hence each of the present volumes has as its core a collection of analytic essays. Some of these have been published or presented elsewhere, but are printed here in revised form. In some cases we have updated the earlier versions in the light of subsequent research, and, where appropriate, have inserted cross-references between contributions. Authorship is in some cases something of a problem, as a number of people have had a hand in collating the data, doing the reconstructions, and (re)writing for publication here. In most chapters, however, one person did the research which determined the structure of the terminology, and that person appears as the first or only author, and where another or others had a substantial part in putting together the paper itself, they appear as the second and further authors.

3 Reconstructing the lexicon

The lexical reconstructions presented in these volumes are arrived at using the standard methods of comparative linguistics, which require as preliminaries a theory of subgrouping (§3.2) and the working out of systematic sound correspondences among cognate vocabulary in contemporary languages (§3.3). As well as cognate sets clearly attributable to POc, we have included some cognate sets which at this stage are attributable to various interstage languages, particularly Proto Western and Proto Eastern Oceanic (but see §3.2 for definitions). We have set out to pay more careful attention to reconstructing the semantics of POc forms than has generally been done in earlier work, treating words not as isolates but as parts of terminologies.

3.1 Terminological reconstruction

Our method of doing ‘terminological reconstruction’ is as follows. First, the terminologies of present-day speakers of Oceanic languages are used as the basis for constructing a hypothesis about the semantic structure of a corresponding POc terminology, taking account of (i) ethnographic evidence, i.e. descriptions of the lifestyles of Oceanic communities and (ii) the geographical and physical resources of particular regions of Oceania. For example, by comparing terms in several languages for parts of an outrigger canoe, or for growth stages of a coconut, one can see which concepts recur and so are likely to have been present in POc. Secondly, a search is made for cognate sets from which forms can be reconstructed to

match each meaning in this hypothesised terminology. The search is not restricted to members of the Oceanic subgroup; if a term found in an Oceanic language proves to have external (non-Oceanic) cognates, the POc antiquity of that term will be confirmed and additional evidence concerning its meaning will be provided. Thirdly, the hypothesised terminology is re-examined to see if it needs modification in the light of the reconstructions. There are cases, highlighted in the various contributions to these volumes, where we were able to reconstruct a term where we did not expect to do so and conversely, often more significantly, where we were unable to reconstruct a term where we had believed we should be able to. In each case, we have discussed the reasons why our expectations were not met and what this may mean for Oceanic culture history.

Blust (1987: 81) distinguishes between conventional ‘semantic reconstruction’, which asks, “What was the probable meaning of protomorpheme X?”, and Dyen and Aberle’s (1974) ‘lexical reconstruction’, where one asks, “What was the protomorpheme which probably meant ‘X’?” At first sight, it might appear that terminological reconstruction is a version of lexical reconstruction. However, there are sharp differences. Lexical reconstruction applies a formal procedure: likely protomeanings are selected from among the glosses of words in available cognate sets, then an algorithm is applied to determine which meaning should be attributed to each set. This procedure may have unsatisfactory results, as Blust points out. Several reconstructions may end up with the same meaning; or no meaning may be reconstructed for a form because none of the glosses of its reflexes is its protomeaning.

Terminological reconstruction is instead similar to the semantic reconstruction approach. In terminological reconstruction the meanings of protomorphemes are not determined in advance. Instead, cognate sets are collected and their meanings are compared with regard to:

- their specific denotations, where these are known;
- the geographic and genetic distribution of these denotations (i.e. are the glosses from which the protogloss is reconstructed well distributed?);
- any derivational relationships to other reconstructions;
- their place within a working hypothesis of the relevant POc terminology (e.g., are terms complementary —‘bow’ implies ‘arrow’; ‘seine net’ implies ‘floats’ and ‘weights’? Are there different levels of classification—generic, specific, and so on?).

For example, it proved possible to reconstruct the following POc terms for tying with cords (vol.1, ch.9, §10):

POc **buku* ‘tie (a knot); fasten’

POc **p^wita* ‘tie by encircling’

POc **paqu(s)*, **paqus-i-* ‘bind, lash; construct (canoe +) by lashing together’

POc **pisi* ‘bind up, tie up, wind round, wrap’

POc **kiti* ‘tie, bind’

In each of the supporting cognate sets from contemporary languages there are a number of items whose glosses in the dictionaries or word lists are too vague to tell the analyst anything about the specific denotation of the item, and in the case of **kiti* this prevents the assignment of a more specific meaning. The verb **buku* can be identified as the generic term for tying a knot because of its derivational relationship (by zero derivation) with a noun whose denotation is clearly generic, **buku* ‘node (as in bamboo or sugarcane); joint; knuckle; knot in wood, string or rope’ (vol.1, ch.4, §3.2). Reconstruction of the meaning of **p^wita* as ‘tie by encircling’ is supported by the meanings of the Lukep, Takia and Longgu reflexes, respectively

‘tie by encircling’, ‘tie on (as grass-skirt)’, and ‘trap an animal’s leg; tie s.t. around ankle or wrist’: Lukep and Takia are North New Guinea languages, whilst Longgu is SE Solomonic. Reconstruction of the meaning of **paku(s)*, **paku-s-i-* as ‘bind, lash; construct (canoe +) by tying together’ is supported by the meanings of the Takia, Kiribati and Samoan reflexes, respectively ‘tie, bind; construct (a canoe)’, ‘construct (canoe, house)’, and ‘make, construct (wooden objects, canoes +)’: Takia is a North New Guinea language, Kiribati is Micronesian, and Samoan is Polynesian. The meaning of **pisi* is similarly reconstructed by reference to the meanings of its Mono-Alu, Mota, Port Sandwich, Nguna and Fijian reflexes.

Often, however, the contributors have often been less fortunate in the information available to them. For example, Osmond (vol.1, ch.8, §9) reconstructs six POc terms broadly glossed as ‘spear’. Multiple terms for implements within one language imply that these items were used extensively and possibly in specialised ways. Can we throw light on these specialised ways? Unfortunately, some of the word lists and dictionaries available give minimal glosses, e.g. ‘spear’, for reflexes of the six reconstructions. What we need to know for each reflex is: what is the level of reference? Is it a term for all spears, or perhaps all pointed projectiles including arrows and darts? Or does it refer to a particular kind of spear? Is it noun or verb or both? If a noun, does it refer to both the instrument and the activity? Most word lists are frustratingly short on detail. For this kind of detail, ethnographies have proved a more fruitful source of information than many word lists.

Another problem is inherent in the dangers of sampling from over 450 languages. The greater the number of languages, the greater are the possible variations in meaning of any given term, and the greater the chances of two languages making the same semantic leaps quite independently. Does our (sometimes quite limited) cognate set provide us with a clear unambiguous gloss, or have we picked up an accidental bias, a secondary or distantly related meaning? Did etymon *x* refer to fishhook or the material from which the fishhook was made? Did etymon *y* refer to the slingshot or to the action of turning round and round?

3.2 Subgrouping and reconstruction

3.2.1 Subgrouping

Although the subgrouping of Austronesian languages and questions about which protolanguage was spoken where remain in some cases somewhat controversial, it is impossible to proceed without making some assumptions about these matters. Figures 1.1 and 1.2 are approximate renderings of our subgrouping assumptions. The upper part of the tree, shown in Figure 1.1, is due to Blust, originally presented in Blust (1977*b*) and repeated with additional supporting evidence in subsequent publications (1978, 1982, 1983–84, 1993).⁷ The diagram of the lower (Oceanic) part of the tree in Figure 1.2 shows nine primary subgroups of Oceanic. Its rake-like structure indicates that no convincing body of shared innovations has been found to allow any of the nine subgroups to be combined into higher-order groupings. Sections 3.2.2, 3.2.3, and 3.2.4 offer some commentary on our subgrouping, and in §3.2.3 we explain how we handle the rake-like structure in making reconstructions.

⁷ For commentaries on Austronesian subgrouping, see Ross (1995) and Adelaar (2005).

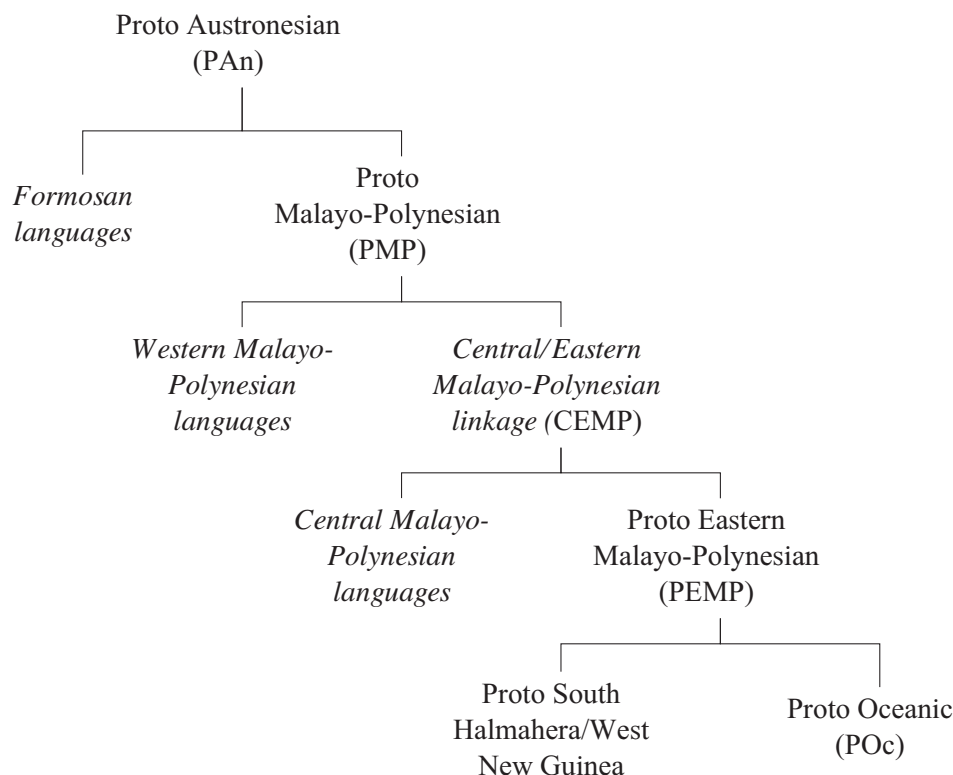


Figure 1.1 Schematic diagram showing higher-order subgroups of Austronesian languages.

3.2.2 Kinds of subgroup

In Figures 1.1 and 1.2 each node is either a single language,⁸ usually a reconstructed protolanguage, or, in italics, a group of languages.

Where a node is a protolanguage, its descendants form a proper subgroup (in the technical sense in which historical linguists use the term ‘subgroup’). A proper subgroup is identified by innovations shared by its member languages, i.e. it is ‘innovation-defined’ in the terminology of Pawley & Ross (1995). These innovations are assumed to have occurred just once in the subgroup’s protolanguage, i.e. the exclusively shared ancestor of its members. Thus languages of the large Oceanic subgroup of Austronesian share a set of innovations relative to the earlier Austronesian stages shown in Figure 1.1 (Dempwolff 1934).⁹ By inference these innovations occurred in their common ancestor, POc, and the claim that they are innovations is based on a comparison of reconstructed POc with reconstructed PMP. The innovations may be phonological (e.g. PMP **e*, pronounced [ə], and PMP **aw* both became POc **o*), morphological (e.g. POc acquired a morphological distinction between three kinds of possessive relationship: food, drink and default), or lexical (e.g. PMP **limaw* ‘citrus fruit’ was replaced by POc **molis*).

⁸ The two very closely related languages Mussau and Tench form a minor exception.

⁹ Chapter 4 of Lynch et al. (2002) gives a recent account of these innovations.

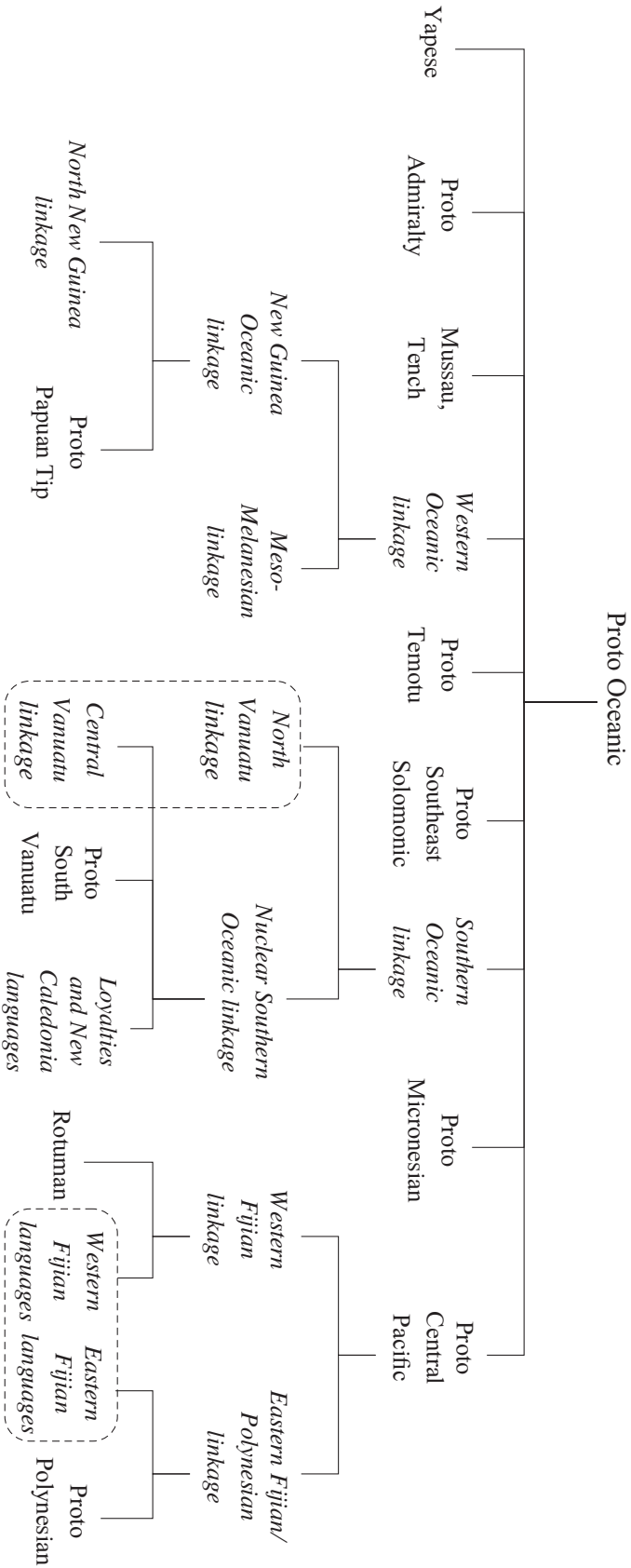


Figure 1.2 Schematic diagram showing the subgrouping of Oceanic Austronesian languages.

Italics are used in Figures 1.1 and 1.2 to indicate a group of languages which is not a proper subgroup, i.e. has no identifiable *exclusively* shared parent. Thus *Formosan languages* in Figure 1.1 indicates a collection of languages descended (along with PMP) from PAN. They are spoken in Taiwan, but do not form a subgroup. There was no ‘Proto Formosan’, as Formosan languages and language groups are all descended directly from PAN.

Some of the italicised labels in Figures 1.1 and 1.2 include the term *linkage*. A linkage (an ‘innovation-linked group’ in the terminology of Pawley & Ross 1995) is a collection of usually quite closely related languages or dialects,¹⁰ speakers of which were in sufficient contact at one time or another during their history for innovations to pass from one language to the next, often resulting in a pattern such that the domains of various innovations overlap but are not coterminous.¹¹ A number of Oceanic linkages were recognised during the 1980s (Geraghty 1983, Pawley & Green 1984, Ross 1988).¹² A linkage may arise in at least three ways, but distinguishing between them is often impossible.

First, what would otherwise be a proper subgroup may happen to lack exclusively shared innovations, perhaps because the parent did not exist as a unit for long enough to undergo any innovations of its own.¹³

Second, a linkage may consist of some but not all of the languages descended from a single parent. The Western Oceanic linkage reflects the innovations of POc, but no innovation is common to the whole of Western Oceanic (although the merger of POc **r* and **R* comes close). However, the languages of its three component linkages—North New Guinea, Papuan Tip and Meso-Melanesian—display complex patterns of overlapping innovations. The Western Oceanic linkage appears to be descended from the dialects of POc that were left behind in the Bismarck Archipelago after speakers of the languages ancestral to the other eight primary subgroups in Figure 1.2 had moved away to the north or east. After these departures various innovations occurred. Each arose somewhere in the Western Oceanic dialect network and spread to neighbouring dialects without reaching every dialect in the network.

The third type of linkage is the result of contact among languages descended from more than one immediate parent. An example is the Fijian linkage, which represents a reintegration of parts of earlier Western and Eastern Fijian linkages (Geraghty & Pawley 1981, Geraghty 1983, Pawley 1996b).¹⁴ Geraghty reconstructed the history of the Fijian linkage by painstaking analysis of innovations from at least two stages in its history. From the earlier period Western Fijian languages share innovations with Rotuman and Eastern Fijian with Polynesian. From a more recent period Western Fijian and Eastern Fijian languages share innovations

¹⁰ In what follows, ‘language’ is used to mean ‘language or dialect’.

¹¹ One or more innovations may spread right across the languages of the linkage. In this case it becomes virtually impossible to distinguish it from a proper subgroup.

¹² Recent work in Indo-European appeals to the concept of linkage: Garrett (2006) suggests that the dialects ancestral to Greek were not dialects of ‘Proto Greek’ but a collection of Nuclear Indo-European dialects drawn together by relations between the communities ancestral to the Greek city states, across which spread the innovations which characterise Ancient Greek.

¹³ A situation in which a subgroup is both proper (i.e. defined by exclusive innovations) and a linkage (displaying overlapping patterns of innovations) is of course possible, the exclusively shared innovations having occurred in the parent, the others after the break-up of the parent. It so happens that we have no need of this construct here.

¹⁴ ‘Eastern Fijian linkage’ in Figure 1.2 is our label for Geraghty’s (1983) ‘Tokalau Fijian’.

with each other, reflecting their reintegration into a single linkage, within which the present Western/Eastern boundary has shifted relative to the (fuzzy) boundary of the earlier period.

For most of the linkages noted in Figures 1.1 and 1.2 this kind of analysis is not available. For example, Blust (1993) indicates that CEMP was a linkage. But its history is far from clear. Was there perhaps a PCEMP that was so shortlived that it underwent no innovations of its own? Or does CEMP perhaps include some languages that share history with languages to their west and others that share history with those to their north? The North/Central Vanuatu linkage, long assumed to be some sort of genealogical unit, appears to reflect the partial reintegration of at least two dialect networks, North Vanuatu and Central Vanuatu, that probably never diverged greatly from each other, but the details of this history are difficult to elucidate (Lynch 2000a).¹⁵

The languages of a linkage have no identifiable exclusively shared parent. Yet we have found many instances in which a cognate set is limited to one of the linkages in Figures 1.1 and 1.2: CEMP, Western Oceanic, New Guinea Oceanic, Southern Oceanic or the reintegrated North and Central Vanuatu linkage. As with PEOc and PROc (§3.2.3), we think it is preferable to attribute these reconstructions to a hypothetical protolanguage rather than to a higher node in the tree. Hence there are reconstructions labelled PCEMP, PWOC and so on. Again these apparent lexical innovations offer only the weakest evidence for the protolanguage to which they are attributed. In addition to the explanations of the kinds offered for PEOc and PROc etyma in §3.2.3 it is possible, for example, that an innovatory ‘PWOC’ etymon arose when the Western Oceanic dialect network was still close-knit, and spread from dialect to dialect before the network broke into the two networks ancestral to its present-day first-order subgroups.

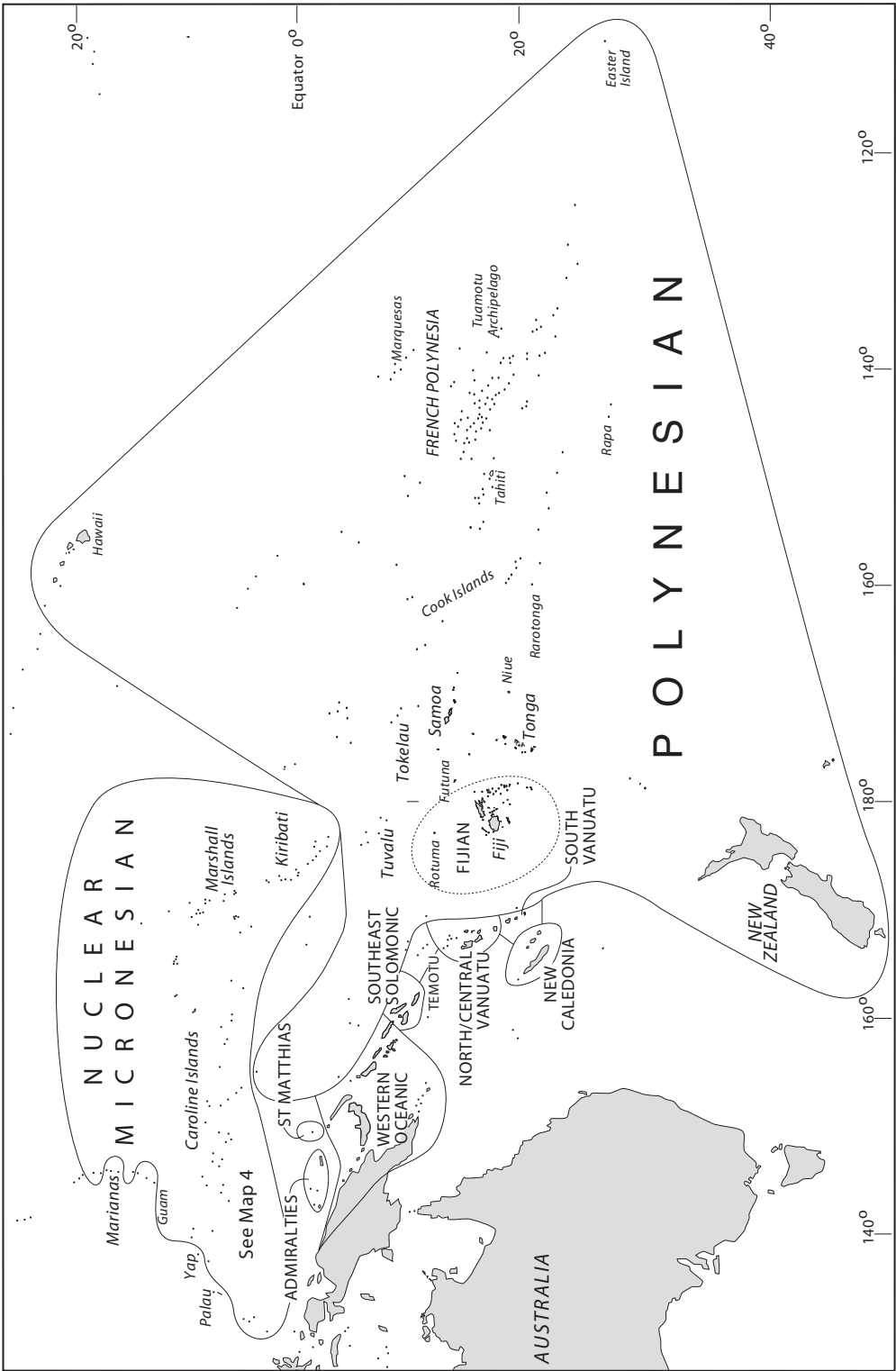
3.2.3 Criteria for reconstruction

The strength of a lexical reconstruction rests crucially on the distribution of the supporting cognate set across subgroups. The *distribution* of cognate forms and agreements in their meanings is much more important than the *number* of cognates. It is enough to make a secure reconstruction if a cognate set occurs in just two languages in a family, with agreement in meaning, provided that the two languages belong to different primary subgroups and provided that there is no reason to suspect that the resemblances are due to borrowing or chance. The PMP term **apij* ‘twins’ is reflected in several western Malayo-Polynesian languages (e.g. Batak *apid* ‘twins, double (fused) banana’) but only a single Oceanic reflex is known, namely Roviana *avisi* ‘twins of the same sex’. Because Roviana belongs to a different first-order branch of Malayo-Polynesian from the western Malayo-Polynesian witnesses and because there is virtually no chance that the agreement is due to borrowing or chance similarity, this distribution is enough to justify the reconstruction of PMP **apij*, POc **apic* ‘twins’.

The rake-like form of Figure 1.2 almost certainly reflects the very rapid settlement of Oceania out of the Bismarcks,¹⁶ but it confronts us with a methodological question. If we follow the rubric that we make a reconstruction if a cognate set occurs in languages of just

¹⁵ For a history of scholarly views of the subgrouping of North and Central Vanuatu languages see Clark (2008: §1.3). For arguments supporting a NCV grouping, see Clark (2008: ch.4).

¹⁶ Approximate dates are as follows. The Bismarcks themselves were settled by around 1400 BC. Archaeology has the Lapita culture in the Temotu Archipelago by about 1200 BC, in New Caledonia by 1100 BC, in Fiji and Tonga by 1000 BC and in Samoa by 900 BC (Green 2003).



Map 3 Groups of Oceanic languages used in cognate sets

two primary subgroups, then reflexes of an etymon in, say, a SE Solomonian language and a Micronesian language would be sufficient evidence for a POc reconstruction and the absence of reflexes in Admiralty and Western Oceanic would be irrelevant. Given what we know about the location of the POc homeland (in the Bismarcks; vol.2, ch.2) and the early eastward spread of Oceanic speakers, this is too loose a criterion. Instead, we assume two hypothetical nodes not shown in the tree in Figure 1.2.¹⁷ These are

- Remote Oceanic, comprising Southern Oceanic, Micronesian and Central Pacific;
- Eastern Oceanic, comprising SE Solomonian and Remote Oceanic.¹⁸

If a cognate set occurs in two or all three of the groups in Remote Oceanic, the reconstruction is attributed to Proto Remote Oceanic (PROc). If a cognate set occurs in one or more of the groups in Remote Oceanic and in SE Solomonian, it is attributed to Proto Eastern Oceanic (PEOc). In this way we acknowledge that such reconstructions may represent an innovation that postdates the spread of the early Oceanic speech community. There are enough PROc and PEOc reconstructions to suggest that such lexical innovations indeed occurred. This in turn provides weak evidence for Remote Oceanic and Eastern Oceanic subgroups, but evidence that is too weak to be relied on, for at least two reasons. First, it is quite possible that some of our PROc and PEOc reconstructions will be promoted to POc as more Admiralty and Western Oceanic data become available. Second, it is reasonable to assume that some of our PROc and PEOc etyma are of POc antiquity but happen to have been lost in Proto Admiralty and Proto Western Oceanic. Without supporting phonological or morphological evidence we are unwilling to treat PROc or PEOc as anything other than convenient hypotheses which allow us to retain rigorous criteria for a POc reconstruction.

In volumes 1 and 2 a reconstruction here labelled ‘PROc’ would have been labelled ‘PEOc’, but the absence of SE Solomonian reflexes from among its reflexes indicates that it has the same status as a PROc reconstruction in the present volume. Two factors have led to the distinction between PEOc and PROc here. One is that, because the primary biogeographic divide in Oceania is between Near and Remote Oceania (see vol. 2, Map 5), i.e. between the Solomon Islands and Vanuatu, the question of whether or not a plant name includes a SE Solomonian reflex is significant, and there are many plant names that do not (and are thus attributed to PROc). The other is that the historical separateness of SE Solomonian from both Western Oceanic and the groups treated as Remote Oceanic has become increasingly clear through recent research (Pawley 2007).

Our criterion for recognising a reconstruction as POc is that the cognate set must occur in at least two out of four criterial groupings: Admiralties (or Yapese or Mussau), Western Oceanic, Temotu and our hypothetical Eastern Oceanic. Both here and at the hypothetical interstages defined above, no reconstruction is made if there are grounds to infer borrowing

¹⁷ We included these nodes in the corresponding tree in Figure 1 of volumes 1 and 2, but this was too easily interpreted as a statement of our views on subgrouping, so we abandon it here and in Appendix 2.

¹⁸ The term ‘Eastern Oceanic’ and the search for evidence of an Eastern Oceanic subgroup has a relatively long pedigree in Oceanic linguistics (Biggs 1965, Pawley 1972, Pawley 1977, Lynch & Tryon 1985, Geraghty 1990). However, by the time volume 1 of the present work was published in 1998 it was already evident that innovations supporting the existence of an Eastern Oceanic subgroup were not forthcoming, as Pawley & Ross (1995: 79) had mentioned in a footnote. Our use of the term here is more inclusive than most, resembling more closely the ‘Central/Eastern Oceanic’ of Lynch & Tryon (1983) (the published version of which, Lynch & Tryon 1985, presents a less inclusive version of Central/Eastern Oceanic) and of Lynch et al. (2002: 94–96), who also express reservations about its status.

from one of these groupings to another.¹⁹ We also reconstruct an etymon to POc if it is reflected in just one of the four criterial groupings and in a non-Oceanic Austronesian language (a member of one of the subgroups on the left branches in Figure 1.1), as illustrated above by the reconstruction of POc **apic* ‘twins’.

These criteria are identical to those applied in volumes 1 and 2 except for the addition of Temotu (which figures in very few cognate sets). The establishment of Temotu as a primary subgroup (Ross & Næss 2007) postdates the publication of volumes 1 and 2. Temotu comprises the languages of the Reef Islands, Santa Cruz, Utupua and Vanikoro, located 400 km east of the main Solomons archipelago and to the north of Vanuatu (Map 3).

There are indications that Yapese (a single-language subgroup) and Mussau and Tench (a subgroup with two closely related languages) may be more closely related to Admiralty than to any other Oceanic subgroup,²⁰ and for this reason they are treated as Admiralty languages for the purposes of reconstruction. That is, the presence of a reflex in one or more of these languages and in Admiralty does not support a POc reconstruction, but the presence of a reflex in one or more of these languages and one of Western Oceanic, Temotu and Eastern Oceanic does support one.

In chapter 2 (§4) of volume 2 Pawley discusses Blust’s (1998) proposal that the primary split in Oceanic divides Admiralty from a subgroup embracing all other Oceanic languages. Pawley dubs the latter ‘Nuclear Oceanic’. If Blust’s subgrouping were accepted, then an etymon which lacked cognates outside Oceanic would need to be reflected both in an Admiralties language and in a non-Admiralties language for a POc reconstruction to be made. Etyma with reflexes in both Western and Eastern Oceanic, but not in the Admiralties, would be reconstructed as Proto Nuclear Oceanic. Under the criteria outlined above, however, we attribute these reconstructions to POc. These criteria were used in volumes 1 and 2, and we have thought it wise to maintain them throughout the volumes of this work. The reader who wishes to single out reconstructions attributable to a putative Proto Nuclear Oceanic (rather than to POc) can easily recognise them, however. They are those POc reconstructions for which (i) there are no Admiralties reflexes, and (ii) there is no higher-order reconstruction (i.e. PEMP, PCEMP, PMP or PAN), since the latter would be based on cognates outside Oceanic.

3.2.4 Further notes on subgroups

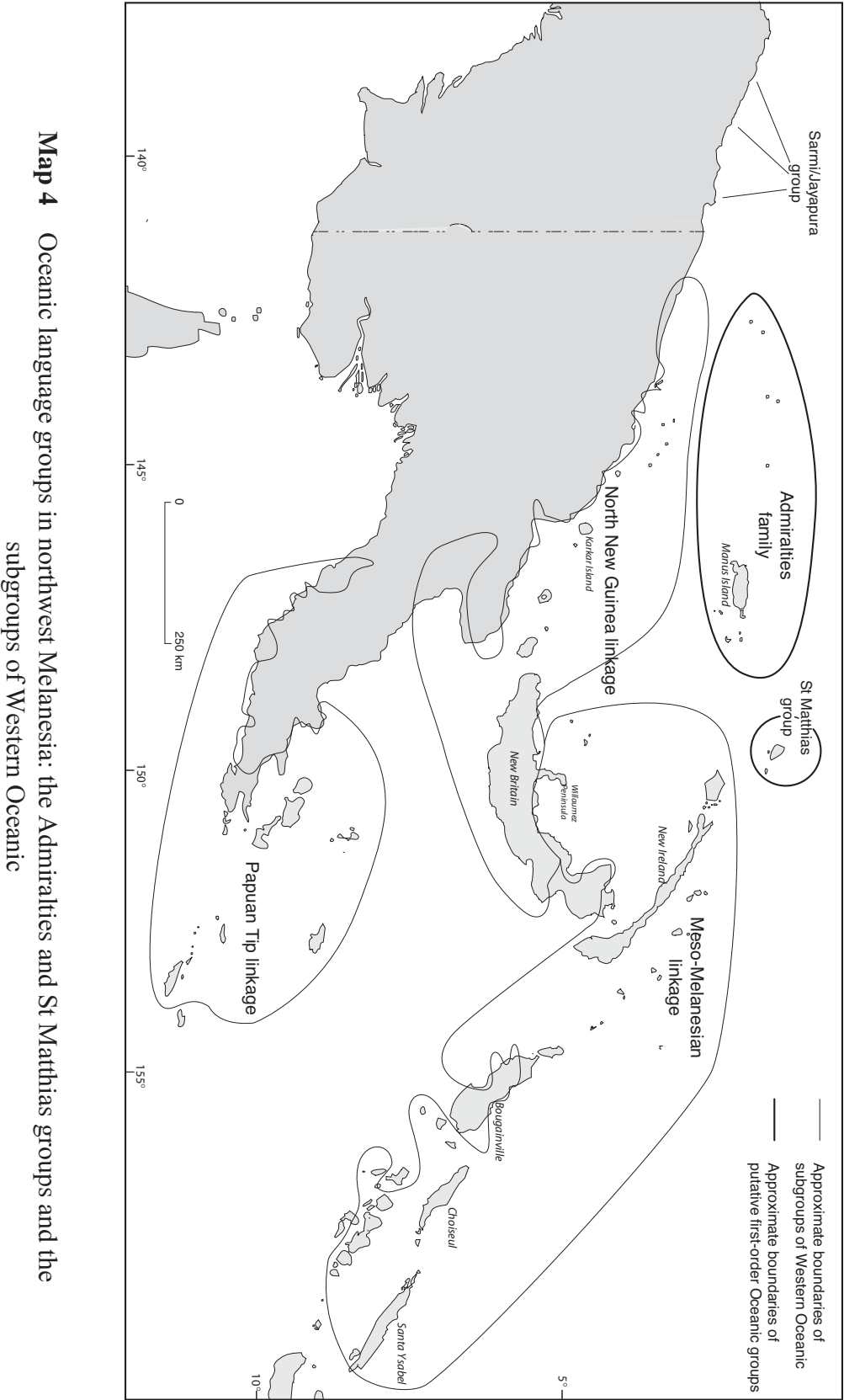
This section brings together brief notes on the subgroups in Figure 1.2 beyond those mentioned in the discussion in §§3.2.3–3.2.2.

Admiralty is a proper subgroup (Ross 1988: ch.9).

Western Oceanic consists of the North New Guinea (NNG), Papuan Tip (PT), Meso-Melanesian (MM) linkages and the Sarimi/Jayapura (SJ) group (see Map 4). The last-named may belong to the NNG linkage, but this is uncertain (Ross 1996*b*). It is not shown in Figure 1.2 and its languages do not play a crucial role in reconstruction. It is likely that the

¹⁹ Cases where such an inference can be made occur mostly at the boundary (in the Solomon Islands) between Western and Eastern Oceanic. Borrowing is likely (and is often reflected in unexpected sound correspondences) where an etymon occurs (i) in Western Oceanic and only in SE Solomonic languages or (ii) in SE Solomonic languages and only in the NW Solomonic languages (a subgroup within the Meso-Melanesian linkage of Western Oceanic).

²⁰ On the positions of Yapese and Mussau, see respectively Ross (1996*a*) and Ross (1988: 315–316, 331).



NNG and PT groups form a super-group, the New Guinea Oceanic linkage, and so etyma reflected only in NNG and PT languages are attributed to a putative Proto New Guinea Oceanic (Milke 1958, Pawley 1978), and etyma reflected in either NNG or PT (or both) and in MM are labelled PWOC.

SE Solomonian was established as a proper subgroup by Pawley (1972: 98–110). Further support was provided by Levy (1979, 1980, n.d.), Tryon & Hackman (1983) and František Lichtenberk (1988).

The *Southern Oceanic* linkage as proposed by Lynch (1999, 2000, 2001, 2004) is characterised by complex overlapping innovations, but by none that are reflected in all its member languages and would qualify it as a proper subgroup (see discussion in Lynch et al. 2002: 112–114).²¹

Micronesian is a proper subgroup (Jackson 1983, 1986, Bender et al. 2003).

Central Pacific is a proper subgroup, but one defined by only a handful of shared innovations, indicating that the period of unity was short (Geraghty 1996). The high-order subgrouping of Central Pacific is due to Geraghty (1983), except for the position of Rotuman, due to Pawley (1996*b*). Within Central Pacific is another long recognised proper subgroup, *Polynesian*, for which Pawley (1996*a*) lists diagnostic innovations.

3.3 Sound correspondences

As we noted above, reconstruction depends on working out the systematic sound correspondences among cognate vocabulary in contemporary languages and on having a working hypothesis about how the sounds of POC have changed and are reflected in modern Oceanic languages. Working out sound correspondences even for twenty languages is a large task, and so we have relied heavily on our own previous work and the work of others. The sound correspondences we have used are those given by Ross (1988) for Western Oceanic and Admiralties; by Levy (1979, 1980) and František Lichtenberk (1988) for Cristobal-Malaitan, by Pawley (1972) and Tryon & Hackman (1983) for SE Solomonian; by Ross & Næss (2007) for Temotu; by Tryon (1976) and Clark (1996, 2008) for North and Central Vanuatu; by Lynch (1978, 2001) for Southern Vanuatu; by Geraghty (1989) and Ozanne-Rivierre (1992) for New Caledonia; by Jackson (1986) and Ross (1996*a*) for Nuclear Micronesian; by Geraghty (1986) for Central Pacific; by Biggs (1978) for Polynesian; by Ross (1996*a*) for Yapese; and by Ross (1996*b*) for Oceanic languages of Irian Jaya.

For non-Oceanic languages we have referred to sound correspondences given by Tsuchida (1976) for Formosan languages; by Zorc (1977, 1986) and Reid (1982) for the Philippines; by Adelaar (1992) and Nothofer (1975) for Malay and Javanese; by Sneddon (1984) for Sulawesi; by Collins (1983) for Central Maluku; and by Blust (1978*a*) for South Halmahera and Irian Jaya.

We are well aware that regular sound correspondences can be interfered with in various ways: by phonetic conditioning that the analyst has not identified (see, e.g., Blust 1996*a*), by borrowing (for an extreme Oceanic case, see Grace 1996), or, as recent research suggests, by the frequency of an item's use (Bybee 1994). We have tried at least to note, and sometimes to account for, irregularities in cognate sets.

²¹ Because it has only been recently proposed, Southern Oceanic does not appear in Figure 1 of volumes 1 and 2.

Table 1.1 Reconstructed paradigm of POC phonemes

<i>*p^w</i>	<i>*p</i>	<i>*t</i>	<i>*c</i>	<i>*k</i>	<i>*q</i>
<i>*b^w</i>	<i>*b</i>	<i>*d</i>	<i>*j</i>	<i>*g</i>	
		<i>*s</i>			
<i>*m^w</i>	<i>*m</i>	<i>*n</i>	<i>*ñ</i>	<i>*ŋ</i>	
		<i>*r</i>			<i>*R</i>
		<i>*dr</i>			
		<i>*l</i>			
<i>*w</i>			<i>*y</i>		
	<i>*i</i>		<i>*u</i>		
	<i>*e</i>		<i>*o</i>		
		<i>*a</i>			

3.4 POC phonology and orthography

Work based on the sound correspondences of both Oceanic and non-Oceanic languages has resulted in the reconstructed paradigm of POC phonemes shown in Table 1.1. The orthography used here and in the POC reconstructions in this work is from Ross (1988), with the addition of **p^w*. POC phonology and its relationship to PMP is discussed in greater detail in Chapter 2 (§2) of volume 1. Since the publication of volume 1, articles by John Lynch have appeared on POC stress (2000) and POC labiovelar phonemes (2002).

Table 1.2 shows two POC orthographies. The first was established by Biggs (1965), for PEOC, and Grace (1969), who applied it to POC. It has been used with a number of variants, separated by a slash in Table 1.2. The second, introduced by Ross (1988), is the one generally used in this work. The terms ‘oral grade’ and ‘nasal grade’ were used by Grace (1969) and have become conventional among Oceanic linguists to refer to the outcomes of certain sound changes that occurred between PMP and POC (vol.1, ch.2, §2.4).

Table 1.2 POC orthographies after Grace (1969) and Ross (1988)

Grace	oral grade	<i>*p</i>	<i>*pw</i>	<i>*t</i>	<i>*d/*r</i>	<i>*s</i>	<i>*j</i>	<i>*k</i>			
Ross		<i>*p</i>	<i>*p^w</i>	<i>*t</i>	<i>*r</i>	<i>*s</i>	<i>*c</i>	<i>*k</i>			
Grace	nasal grade	<i>*mp</i>	<i>*ŋp/*mpw</i>	<i>*nt</i>	<i>*nd/*nr</i>	<i>*nj</i>	<i>*ŋk</i>				
Ross		<i>*b</i>	<i>*b^w</i>	<i>*d</i>	<i>*dr</i>	<i>*j</i>	<i>*g</i>				
Grace		<i>*m</i>	<i>*ŋm/*mw</i>	<i>*n</i>	<i>*ñ</i>	<i>*ŋ</i>	<i>*w</i>	<i>*y</i>	<i>*l</i>	<i>*q</i>	<i>*R</i>
Ross		<i>*m</i>	<i>*m^w</i>	<i>*n</i>	<i>*ñ</i>	<i>*ŋ</i>	<i>*w</i>	<i>*y</i>	<i>*l</i>	<i>*q</i>	<i>*R</i>
Grace		<i>*i</i>	<i>*o</i>	<i>*e</i>	<i>*a</i>	<i>*u</i>					
Ross		<i>*i</i>	<i>*o</i>	<i>*e</i>	<i>*a</i>	<i>*u</i>					

4 Conventions common to the series

4.1 Chapter format

Each of the contributions to these volumes concerns a particular POc ‘terminology’. Generally, each contribution begins with an introduction to the issues raised by the reconstruction of its particular terminology, and the bulk of each contribution consists of reconstructed etyma with supporting data and a commentary on matters of meaning and form.

The reconstruction of POc **kayu* ‘tree’ below, adapted from Chapter 3, §4.2, shows how reconstructions and supporting cognate sets are presented. Above it is Dempwolff’s earlier superordinate (PMP) reconstruction. Below it are supporting reflexes. Contributors vary in the degree to which they insert lower-order reconstructions like PPn **kau* below. Lower-order reconstructions are sometimes given to clarify the relationship of reflexes to the higher-order reconstruction: Southern Vanuatu languages, for example, have undergone so much phonological change that a Proto Southern Vanuatu reconstruction helps explicate the relationship between Southern Vanuatu reflexes and the POc reconstruction. In the set below the Proto Polynesian (PPn) reconstruction shows an extension in its meaning.

PMP **kayu* ‘tree, wood, timber’ (Dempwolff 1938)

POc **kayu* ‘tree or shrub: generic name for plants with woody stems and branches, probably not including palms or tree-ferns; wood, stick’

Adm:	Loniū	<i>ke</i>	‘tree, wood’
Adm:	Titan	<i>kei</i>	‘firewood’
NNG:	Lukep (Pono)	<i>kai</i>	‘tree, wood’
NNG:	Takia	<i>ai</i>	‘tree (generic), wood, firewood, plant’
PT:	Iduna	<i>ai</i>	‘tree, plant, wood, fire, light’
PT:	Motu	<i>au</i>	‘tree, firewood’
MM:	Patpatar	<i>ai-</i>	‘tree species, followed by name of species’
MM:	Mono	<i>au</i>	‘tree’
SES:	Bugotu	<i>yai-</i>	‘tree, shaft of spear’
SES:	Kwaio	<i>?ai</i>	‘branch, tree, stick; woody plant (shrub, tree)’
NCV:	NE Ambae	<i>kai</i>	‘tree, wood’
NCV:	Nguna	<i>na-kau</i>	‘tree’
SV:	Anejoñ	<i>in-yai</i>	‘tree, wood, often used for relatively small bushes’
NCal:	Xârâcùù	<i>k"ãã</i>	‘wood, tree (general term)’
Mic:	Kiribati	<i>kai</i>	‘wood (in general), tree, plant, stick’
Mic:	Ulithian	<i>-xæy</i>	‘counting classifier for trees’
Fij:	Kadavu	<i>kaðu</i>	‘tree, piece of wood, stick’
Fij:	Wayan	<i>kai</i>	‘wood; generic for trees and shrubs, and occasionally also low bushy plants; used in certain compounds as generic for all plants; piece of wood, stick’

PPn **kau* ‘wood, timber, stalk, stem, handle’ (POLLEX)

Pn:	Tongan	<i>kau</i>	‘stalk, stem’
Pn:	Hawaiian	<i>?au</i>	‘handle, staff, stem, bone of lower arm or leg’

Because our supporting data are drawn from such a wide range of languages, the convention is adopted of prefixing each language name with the abbreviation for the genealogical or geographic group to which the language belongs, so that the distribution of a cognate set is more immediately obvious. Table 1.3 is a key to the labels. Figure §1.2 shows the positions of these groups in the Oceanic tree. We have sought to be consistent in always listing these groups in the same order, but contributors vary in the ordering of languages within groups.

Lynch's recent research on Southern Oceanic (§3.2.4) renders the NCV group mildly anomalous, although there is no doubt that it is an integrated dialect network. There are a number of etyma whose reflexes are confined to North and Central Vanuatu, and so we continue to make 'Proto North/Central Vanuatu' reconstructions, even though these perhaps represent a Southern Oceanic term that has been lost in southern Vanuatu and New Caledonia. Where the distribution of reflexes requires it, the chapters in this volume include reconstructions for PROc and for PSOc. Etyma with these distributions were attributed to PEOc in volumes 1 and 2, but the distributions are transparent, thanks to the presence of the group labels in cognate sets.

In the interests of space we have not given the history of the reconstructions themselves, as this would often require commentary on the modifications made by others and by us, and on why we have made them. Where a reconstruction is not new, we have tried to give its earliest source, e.g. 'Dempwolff 1938' above, but this is difficult when earlier reconstructions differ in form and meaning.

In general, the contributions to these volumes are concerned with items reconstructable in POc, PWOc, PEOc, PROc and occasionally Proto New Guinea Oceanic (PNGOc). Etyma for PWOc, PNGOc and PEOc are reconstructed because these may well also be POc etyma for which known reflexes are not well distributed (see discussion in §3.2.3). The contributors vary in the degree to which they reconstruct etyma for interstages further down the tree.

Table 1.3 Abbreviations for the genealogical or geographic groups

Yap:	Yapese (one language)
Adm:	Admiralty and Mussau/Tench
SJ:	Sarmi/Jayapura
NNG:	North New Guinea
PT:	Papuan Tip
MM:	Meso-Melanesian
SES:	Southeast Solomonian
TM:	Temotu
NCV:	North/Central Vanuatu, i.e. the reintegrated network formed by the North and Central Vanuatu linkages
SV:	Southern Vanuatu
NCal:	Loyalty and New Caledonia
Mic:	Micronesian
Fij:	Fijian, i.e. the reintegrated network formed by Western and Eastern Fijian dialects
Pn:	Polynesian

Reconstructions for lower-order interstages are decreasingly likely to reflect POc etyma and may be the results of cultural change as Oceanic speakers moved further out into the Pacific.

Contributors have usually not sought to make fresh reconstructions at interstages super-ordinate to POc. What they have done, however, is to cite other scholars' reconstructions for higher-order interstages, as these represent a summary of the non-Oceanic evidence in support of a given POc reconstruction. These interstages are shown in Figure 1.1, together with their abbreviations.

Sometimes non-Oceanic evidence has been found to support a POc reconstruction where no reconstruction at a higher-level interstage has previously been made. In this case a new higher-order reconstruction is made, and the non-Oceanic evidence is given in a footnote.

Whilst we have tried to use the internal organisation of the lexicons of Oceanic languages themselves as a guide in setting the boundaries of each terminology, we have inevitably taken decisions which differ from those that others might have made. There are, obviously, overlaps and connections between various semantic domains and therefore between the contributions here. We have done our best to provide cross-references, but we have sometimes duplicated information rather than ask the reader repeatedly to look elsewhere in the book. Indexes at the end of each volume and in the final volume are intended to make it easier to use the volumes collectively as a work of reference.

4.2 Data

Data sources are listed in Appendix 1.

For some reconstructed etyma only a representative sample of reflexes is given. We have endeavoured to ensure, however, that in each case this sample not only is geographically and genetically representative, but also provides evidence to justify the shape of the reconstruction. Where only a few reflexes are known to us, this is usually noted.

Although there are accepted or standard orthographies for a number of the languages from which data are cited here, all data are transcribed into a standard orthography based on that used by Ross (1988: 3–4) in order to facilitate comparison.²² This means, for example, that the *j* of the German-based orthographies of Yabem and Gedaged becomes *y*, Yabem *c* becomes *ʔ*, Gedaged *z* becomes *ʈ* and so on; the *ng* of English-based orthographies becomes *ŋ*; and Fijian *g*, *q* and *c* become *ŋ*, *g* and *ð* respectively.

Unless otherwise indicated, the following symbols have their usual phonetic values: *ð*, *g*, *ɠ*, *ɣ*, *h*, *k*, *l*, *ʈ*, *ʡ*, *m*, *n*, *ŋ*, *ɲ*, *p*, *q*, *χ*, *ɾ*, *r*, *s*, *t*, *w*, *x*, *z*, *ʔ*, *a*, *æ*, *e*, *ɛ*, *ə*, *i*, *ɨ*, *o*, *œ*, *ɔ*, *ʌ*, *u*, *ʊ*. The voiced stops *b*, *d*, *g* and the voiced bilabial trill *ʙ* are prenasalised in some languages, but prenasalisation is not written unless it is phonemically distinctive. Other orthographic symbols (with values in IPA) are:

<i>f</i>	[ɸ, f]	voiceless bilabial or (less often) labio-dental fricative
<i>v</i>	[β, v]	voiced bilabial or (less often) labio-dental fricative
<i>c</i>	[ts], [tʃ]	voiceless alveolar or palatal affricate
<i>j</i>	[dʒ], [dʒ]	voiced alveolar or palatal affricate
<i>y</i>	[j]	palatal glide
<i>dr</i>	[ʳ]	prenasalised voiced alveolar trill (as in Fijian)

²² The main reason for retaining Ross' orthography was that the electronic files initially used in this project were drawn in large part from those used in the research reported in Ross (1988).

ö	[ø]	rounded mid front vowel
ü	[y]	rounded high front vowel

Other superscripts and diacritics are as follows:

- contrastive long vowels are represented by a macron, e.g. *ā*;
- contrastive vowel nasalisation in New Caledonian languages is represented by a circumflex, e.g. *â*;
- labialisation is marked by a superscript *w*, e.g. *p^w*;
- velarisation is marked by a superscript *u*, e.g. *p^u*;
- contrastive aspiration is marked by a superscript *h*, e.g. *p^h*;
- apicolabials are represented by the corresponding apical symbol and the linguolabial diacritic (the ‘seagull’), e.g. *ɬ̥̟*;
- retroflexes are represented by the corresponding apical symbol with a dot beneath, e.g. *ɻ̣*.

Except for inflexional morphemes, non-cognate portions of reflexes, i.e. derivational morphemes and non-cognate parts of compounds, are shown in parentheses (...). Where an inflexional morpheme is an affix or clitic and can readily be omitted, its omission is indicated by a hyphen at the beginning or end of the base. This applies particularly to possessor suffixes on directly possessed nouns (vol.1, ch.2, §3.2). Where an inflexional morpheme cannot readily be omitted, then it is separated from its base by a hyphen. This may happen because of complicated morphophonemics or because the morpheme is always present, like the adjectival *-n* in some NNG and Admiralties languages and prefixed reflexes of the POc article **na* in scattered languages. When a reflex is itself polymorphemic (i.e. the morphemes reflect morphemes present in the reconstructed etymon) or contains a reduplication, the morphemes or reduplicates are also separated by a hyphen.

Languages from which data are cited in this volume are listed in Appendix B in their subgroups, together with an index allowing the reader to find the subgroup to which a given language belongs. Appendix B also includes alternative language names. The difficulty of deciding where the borderline between dialect and language lies, combined with the fact that these volumes contain work by a number of contributors, has resulted in some inconsistency in the naming of dialects in the cognate sets (cf the cognate set supporting POc **kayu* on ch.3, §4.2). Some occur in the form ‘Kara (E)’, i.e. the East dialect of the Kara language, or ‘Lukep (Pono)’, i.e. the Pono dialect of the Lukep language, whilst others are represented simply by the dialect name, e.g. Iduna, noted in Appendix B as ‘Iduna (= dialect of Bwaidoga)’.

4.3 Conventions used in representing reconstructions

Reconstructions are marked with an asterisk, e.g. **kayu* ‘tree’, in keeping with the standard convention in historical linguistics. POc reconstructions, and also PWOC, PEOc and PNGOC reconstructions, are given in the orthography of §3.4. For reconstructions at higher-order interstages the orthographies are those used by Blust in his various publications and the ACD. Reconstructions at lower-order interstages are given in the standard orthography adopted for data (§4.2). Geraghty’s (1986) PCP orthography, for example, is based on Standard Fijian spelling, and is converted into our standard orthography in the same way as Fijian spelling is. In practice, this means that the orthographies for PROc and PCP are the same as for POc, except that a distinction between **p* and **v* is recognised and **R* is absent. Biggs and Clark’s PPn

Table 1.4 Bracketing and segmentation conventions in protoforms

(x)	it cannot be determined whether x was present
(x,y)	either x or y was present
$[x]$	the item is reconstructable in two forms, one with and one without x
$[x,y]$	the item is reconstructable in two forms, one with x and one with y
$x-y$	x and y are separate morphemes
$x-$	x takes an enclitic or a suffix
$\langle x \rangle$	x is an infix

reconstructions are in any case written in an orthography identical to our standard. Bracketing and segmentation conventions in protoforms are shown in Table 1.4.

PMP final consonants are usually retained in POc. However, it happens fairly often that the final consonant in a higher-order reconstructed etymon (e.g. $*-R$ in PMP $*kamaliR$ ‘men’s house’) is not evidenced in any Oceanic reflex because POc final consonants are regularly lost in all the daughter languages from which reflexes are drawn, and we therefore have no evidence as to whether or not the final consonant was retained in the POc etymon in question. In such cases the consonant is reconstructed in brackets (e.g. POc $*kamali(R)$).

In presenting words that display anomalies of form, it is often necessary to posit an expected form. For example, in ch.2, §5, the Gela form *ao* is presented in support of PEOc $*wao$ ‘forest’. Given the reconstruction, however, we would expect the Gela form to be *wao*. In this volume we use a less widely employed convention and mark expected forms with a dagger, e.g. †*wao*, to distinguish them both from reconstructions and real data.²³

There are occasions on which we need to posit a hypothetical form in a reconstructed protolanguage. In such cases the dagger and asterisk conventions are used together. For example in ch.4, §2.4, POc $*lali(c,t)$ ‘buttress roots’ is reconstructed. This is a reflex of PMP $*dali$, but the first consonant has undergone assimilation to the second in POc: the expected (but unattested) POc form would be † $*ralic$.

When historical linguists compile cognate sets they commonly retain word for word the glosses given in the sources from which the items are taken. However, again in the interests of standardisation, we have often reworded (and sometimes abbreviated) the glosses of our sources, while preserving the meaning. Where glosses were in a language other than English we have translated them. In the interests of space and legibility, and because data often have multiple sources, we have given the source of a reflex only when it is not included in the listings in Appendix A.

In glosses we use the conventional abbreviations ‘k.o.’ (as in ‘k.o. yam’) for ‘kind of’, ‘s.o.’ for ‘someone’, ‘s.t.’ for ‘something’, ‘sp.’ for a species and ‘spp.’ for more than one species.

In putting together cognate sets we have quite often found apparent reflexes which do not quite ‘fit’ the set: either they display an unexplained phonological irregularity or their meaning is just a little too different from the rest of the set for us to assume cognacy. Rather

²³ Another convention sometimes used for this purpose is a double asterisk, e.g. $**wao$: we prefer the dagger on aesthetic grounds.

than eliminate them our authors often include them below the cognate set under the rubric ‘cf. also’.

5 Data sources and conventions specific to this volume

5.1 Cognate sets and reconstructions

As explained in §4.1, a cognate set is almost always headed by a reconstruction for the highest-order Oceanic protolanguage that the set allows us to infer. In this volume these reconstructions are mostly for Proto Oceanic, Proto Western Oceanic, Proto Eastern Oceanic or Proto Remote Oceanic, and less often for Proto Southern Oceanic or Proto Central Pacific. Since Oceanic speakers spread across the Pacific from the Bismarck Archipelago to Fiji with remarkable speed, reconstructions at any of these interstages are likely to be early. Proto Admiralty, Proto New Guinea Oceanic and Proto Meso-Melanesian reconstructions are also listed in the few cases where they occur, as there is a fair probability that they are also of early Oceanic antiquity. ‘Proto North/Central Vanuatu’ reconstructions are given where reflexes are restricted to North/Central Vanuatu, even though it is improbable that these languages form an exclusive subgroup (§3.2.2). A PNCV reconstruction is more likely to represent a PSOc term that has not survived in southern Vanuatu and New Caledonia.

Proto Polynesian and Proto Micronesian terms that are not reconstructable to an earlier interstage are not given here, partly because they are likely to represent significantly later developments than the interstages enumerated in the previous paragraph, and partly because they are available elsewhere (Proto Polynesian in POLLEX and Proto Micronesian in Bender et al. 2003).

Lower-order reconstructions are also given *within* a cognate set in certain circumstances. If a term has undergone formal change or has been reapplied to a different plant at a particular point in its history, then a reconstruction is given for the relevant interstage. Also often included are Proto Southern Vanuatu reconstructions from Lynch (2001*c*, 2004*a*), as the phonological histories of Southern Vanuatu languages are far from transparent and the reconstruction often helps to illuminate the cognacy of SV items.

5.2 Botanical and ethnographic information and vernacular names

Sources of the Oceanic plant-name data used to make the reconstructions in this volume are listed in Appendix 1 after general data sources. Four works which provide Austronesian plant names in non-Oceanic languages have been of particular use in the search for cognates which allow the reconstruction of higher-order (PCEMP and PMP) plant names. These works are:

- *A dictionary of Philippine plant names* by Domingo A. Madulid (= Madulid 2001*a*, Madulid 2001*b*);
- *De nuttige planten van Indonesië* [The useful plants of Indonesia] by K. Heyne (= Heyne 1950): this is a catalogue of the useful plants of Indonesia, and includes indigenous plant names from various Indonesian languages of Indonesia;
- *Dictionary of Manggarai plant names* by Jilis A.J. Verheijen (= Verheijen 1982);
- *Dictionary of plant names in the Lesser Sunda Islands* by Jilis A.J. Verheijen (= Verheijen 1990).

Certain works which catalogue plants and give information about their forms, habitats and uses have been consulted over and over again, as will be obvious from the recurrence of certain references. They are:

- *Flora of the Bismarck Archipelago for naturalists* by P.G. Peekel (= Peekel 1984), translated by E.E. Henty from the unpublished German manuscript *Illustrierte Flora des Bismarck-Archipels für Naturfreunde*, compiled by Peekel in New Ireland during the 1930s and completed at Vunapope (Gazelle Peninsula, New Britain) in 1947.
- *A guide to the useful plants of Solomon Islands* by C.P. Henderson and I.R. Hancock (= Henderson & Hancock 1988)
- *Na masu'u 'i Kwara'ae—Our forest of Kwara'ae* by Michael Kwa'ioloa and Ben Burt (= Kwa'ioloa & Burt 2001)
- *Kiladi oro vivineidi ria tingitonga pa idere oro pa goana pa Marovo—Reef and rainforest: An environmental encyclopaedia of Marovo Lagoon, Solomon Islands* by Edvard Hviding (= Hviding 2005)
- *Some common trees of the New Hebrides and their vernacular names* by Sheila Gowers (= Gowers 1976)
- *A guide to the common trees of Vanuatu: With lists of their traditional uses and ni-Vanuatu names* by J.I. Wheatley (= Wheatley 1992)

All six works contain vernacular plant names. Since the Bismarck Archipelago is central to the study of POc plant names, it is a pity that there is no modern work dealing with the plants of the region parallel to, say Henderson & Hancock (1988) for the Solomons or Wheatley (1992) for Vanuatu. The *Guide to the trees of Papua New Guinea* website (Conn & Damas 2006) has been helpful, but still has many gaps in its coverage of the Bismarck Archipelago. Peekel's coverage is excellent and includes all plants (even seagrasses). Indeed his coverage is considerably better than that of either Henderson & Hancock (1988), which covers 'useful plants', or Wheatley (1992), which is restricted to trees. The ethnobotanical works on the Solomons, Kwa'ioloa & Burt (2001) and Hviding (2005), also provide good coverage, but this is balanced by rather thin information about many plants.

Pacific food plants have a literature of their own, and here three works were extensively consulted in addition to those listed above:

- *Les plantes alimentaires de l' Océanie : origines, distribution et usages* [Food plants of Oceania: origins, distribution and uses] by Jacques Barrau (= Barrau 1962)
- *Food plants of Papua New Guinea* by Bruce R. French (= French 1986)
- *Fruits of Oceania* by Annie Walter and Chanel Sam (= Walter & Sam 2002)

Of these only the last provides vernacular terms, with a strong bias towards Vanuatu.

5.3 Scientific plant names

The scientific names of plants are a minefield for the non-botanist, as there are often synonyms and the accepted name for a plant may change over time. In order to keep track of synonyms frequent use was made of the *International plant names index* (IPNI = IPNI 2004) and the *Australian plant names index* (APNI = APNI 1991).

Linnaean families have undergone recent changes. For present purposes the most important is that Leguminosae are subdivided into Cisalpiniaceae (formerly Cisalpinioideae), Fabaceae (formerly Papilionatae) and Mimosaceae (formerly Mimosoideae). Other changes relevant here are that Gramineae (grasses) are now Poaceae, and Palmae are now Arecaceae.

5.4 English and Pacific pidgin plant names

Wherever plant names in English and/or one of the three major Pacific pidgins are available, they are included in the section titles in chapters 5 to 13. Names in the three pidgins are marked with an abbreviation. Together with major sources of names, these are:

TP: Tok Pisin (New Guinea pidgin), Mihalic (1971), French (1986)

P: Pijin (Solomons pidgin), Jourdan (2002)

B: Bislama (Vanuatu pidgin), Gowers (1976), Wheatley (1992), T. Crowley (1995)

The pidgin names are given to help some readers to recognise the plant under discussion. In Papua New Guinea it is common for both indigenous people and expatriates to know a plant by its Tok Pisin name but to know neither an English nor a scientific name. The situation is presumably similar in the Solomon Islands and Vanuatu.

5.5 Indexes

The volume has four indexes. The first, as in previous volumes, is an index of reconstructions arranged by their protolanguages. The second is an alphabetical list of reconstructions, the third an index of plants by genus and species, the fourth by botanical family.