

# Papuan pasts: cultural, linguistic and biological histories of Papuan-speaking peoples

edited by  
Andrew Pawley, Robert Attenborough,  
Jack Golson and Robin Hide



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 2005

National Library of Australia Cataloguing-in-Publication entry:

Papuan pasts: cultural, linguistic and biological  
histories of Papuan-speaking peoples

Bibliography.  
ISBN 0 85883 562 2

1. Papuans - History. 2. Papuans - Social life and customs.  
3. Papuan languages. I. Pawley, Andrew. II. Australian National University.  
Research School of Pacific and Asian Studies. Pacific Linguistics.

995.3

**Cover:** Bird-shaped pestle from Wonia, Western Province, Papua New Guinea  
(AM Reg E065762).

Photographer: John Fields, copyright Australian Museum, Sydney.

Copied and typeset by Margaret Forster  
Cover design by Addcolour Digital  
Printed and bound by Addcolour Digital, Fyshwick, Canberra

# 3 *The chequered career of the Trans New Guinea hypothesis: recent research and its implications*

---

ANDREW PAWLEY

## 1 Introduction

The Trans New Guinea (TNG)<sup>1</sup> hypothesis, which posits a common ancestry for over 400 of the non-Austronesian languages of Melanesia and eastern Indonesia, has had a chequered career.<sup>2</sup> It was a promising child whose birth and early growth were attended by much excitement. There followed a period when it was rejected, or at least regarded with great suspicion by most of the small band of historical linguists knowledgeable about Papuan languages. In the last few years evidence has been brought forward that strongly supports a modified version of the hypothesis.

The main discussion begins by sketching the rather complex history of the hypothesis (section 2). The first thing to note is that there is not one Trans New Guinea hypothesis but a number of variant proposals, each significantly different in scope, reflecting changes in knowledge and interpretation over the past 35 years. Four main variants will be discussed. I will assess the arguments advanced in the 1970s to support the early versions of the TNG hypothesis and will outline the reasons why the critics were unwilling to accept these arguments. In section 3 I turn to the different lines of argument and evidence that have formed the basis of a revised form of the hypothesis. Section 4 will briefly review current views of the internal relationships of the languages assigned to the TNG family. The chapter will end with a consideration of several thorny questions: Where was proto TNG spoken and when did it break up? What circumstances enabled the TNG languages to spread over many areas of New Guinea while preventing them from occupying others?

---

<sup>1</sup> Abbreviation used in this paper are: ANU – Australian National University; bp, BP – (years) before the present; C – consonant; CSNG – Central and South New Guinea; D – dual; IE – Indo-European; ka – thousand years; NENG – North-East New Guinea; P – plural, Proto; pl. – plural; PNG – Papua New Guinea; S – singular; sing. – singular; TNG – Trans New Guinea; V – vowel; 1 – first person; 2 – second person; 3 – third person.

<sup>2</sup> I am indebted to Meredith Osmond for research assistance, carried out with the support of ARC grants to the University of Sydney for Papuan Comparative and Descriptive Studies.

## 2 The history of the hypothesis

### 2.1 On the precursors of the Trans New Guinea hypothesis

It is customary to divide the indigenous languages of the south-west Pacific into three categories: Austronesian, Papuan and Australian. Austronesian is a well-defined family. The same cannot be said of the other two. Australian Aboriginal languages, numbering perhaps 250 when Europeans colonised Australia, have not been demonstrated to be a single family (Bower & Koch 2003; Dixon 1997, 2002; Evans 2003). (I use ‘family’ to refer to the largest established genetic group to which a language belongs.) However, it is generally held that most Australian languages outside of a region of north-west Australia centred in Arnhem Land and the Kimberleys are related. The term ‘Papuan’ is no more than a convenient name for a residual category (Foley 1986). Into the ‘Papuan’ basket are thrown the 700–800 indigenous languages spoken in Papua New Guinea, Indonesia and the Solomon Islands which are not Austronesian. On present evidence these languages divide into more than 20 families and a dozen or so isolates (Foley 2000; Ross 2000, 2001, this volume).

Until the last decades of the 19th century the languages of New Guinea were almost completely unknown to linguists. The imposition of European colonial administrations in that era initiated a period of linguistic research, at first mainly carried out by missionary scholars. In 1893 the English linguist S.H. Ray observed that some of the languages of the New Guinea area do not belong to the Austronesian family. Over the next 60 years, as western exploration of the interior of New Guinea and other large islands proceeded, it became evident that there were hundreds of such languages and that they were genetically and structurally extremely diverse. Before the 1940s no sizeable genetic groups had been identified among the ‘Papuan’ languages—none with more than about 20 members. The situation has changed a good deal since then. In the late 1940s Arthur Capell, from the University of Sydney, showed that at least some of the languages of the central highlands of the then Trust Territory of New Guinea were almost certainly related because they share a number of specific lexical and grammatical resemblances (Capell 1948–1949). Ten years later Stephen Wurm, soon after taking up a post at the Australian National University (ANU), set out to build on Capell’s groundwork. During a nine-month survey of the central Highlands in 1958–1959 he obtained quite extensive data on over 50 languages. In a series of papers (especially Wurm 1960, 1964, 1965, 1971) he argued for a genetic grouping of some 50 languages which he called the East New Guinea Highlands stock. His grounds for this proposal were mainly: (1) lexicostatistical percentages, (2) typological features that he considered diagnostic. Although these were not the kinds of evidence used in classical demonstrations of genetic relationship, the quantity of the arguments he advanced was impressive and there is little doubt that he was right in his main conclusion.

During the 1960s Wurm built up a research team at the ANU whose main agenda was to make sense of the relationships among the Papuan languages. This team, consisting of his colleagues (largely drawn from his own graduate students) and PhD students, set about working in different regions of Melanesia, but mainly in Papua New Guinea.<sup>3</sup>

---

<sup>3</sup> In addition to the staff members of the Department of Linguistics then working in New Guinea—Wurm, Dutton, Voorhoeve and Laycock—several PhD students who had lengthy field experience in PNG made key contributions, especially Karl Franklin, Alan Healey, Ken McElhanon (all members of the Summer Institute of Linguistics) and John Z’graggen SVD.

In the 1960s other scholars besides the ANU group put forward ambitious proposals concerning the relationships of Papuan languages. In a paper drafted much earlier but not published until 1971 that arch ‘lumper’, Joseph Greenberg, suggested that all the Papuan languages of Melanesia and of Timor and Alor belong to a vast ‘Indo-Pacific’ group, to which he also assigned the Andaman Islands and Tasmanian languages. The languages of mainland Australia were excluded. Greenberg’s Indo-Pacific proposal rested mainly on a tenuous chain of resemblances in lexical forms (84 sets) and grammatical forms (10 sets). The resemblances were tenuous because of the uneven distribution of forms across language groups and the lack of means to distinguish chance and borrowing from shared retention. The paper drew almost no response from specialists. This lack of response, I believe, represented two things: (a) extreme scepticism, and (b) the difficulty of disproving a claim of this kind until linguists have established a core of well-defined genetic groups among the languages concerned and have worked out the essentials of their historical development. Within Indo-Pacific, Greenberg posited various smaller groups. He divided the non-Austronesian languages of New Guinea Papuan into seven major groups. One of these, his ‘Central’ group, included all the central highlands languages from the Baliem Valley in West Papua to the Huon Peninsula group in Morobe Province, Papua New Guinea. Evidence for such a group was however not given except as part of the etymologies adduced in support of Indo-Pacific as a whole. In their survey of world languages Carl and Florence Voegelin (1977) also proposed several large groupings of Papuan languages. However, the Voegelins gave no new evidence, being content to briefly summarise the proposals of Greenberg, Wurm and various others.

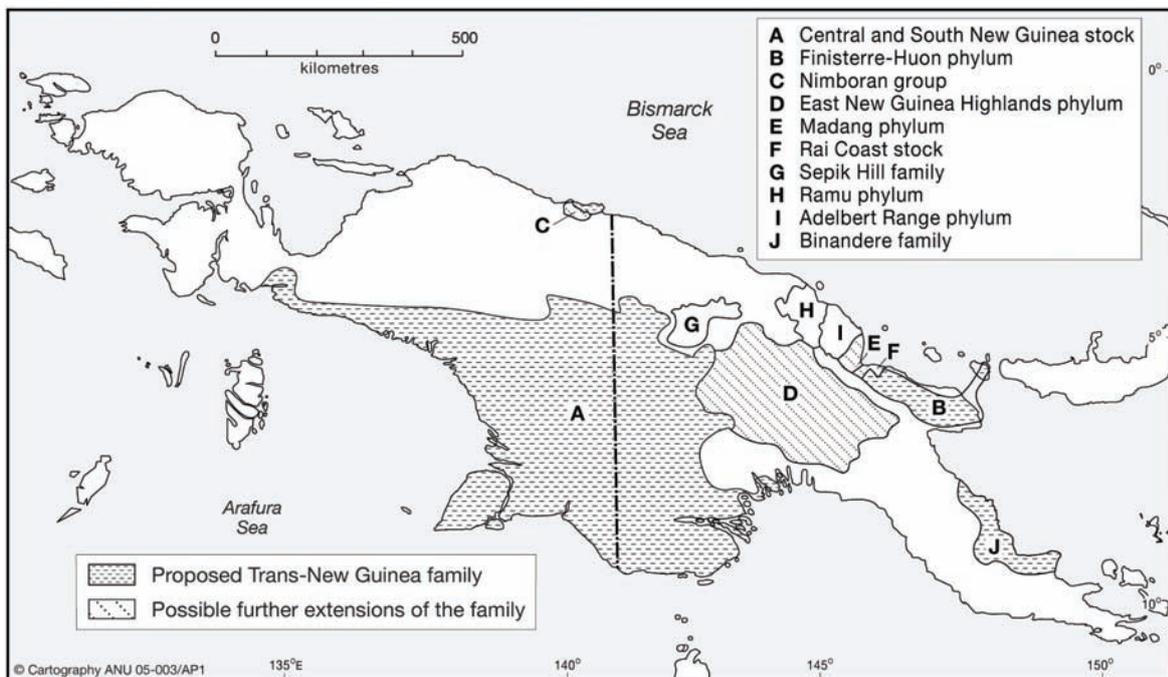
## **2.2 Trans New Guinea I**

The first formulation of a hypothesis about a family given the label ‘Trans New Guinea’ (TNG) emerged from dialogue between two scholars working in widely separated regions of New Guinea. In the mid 1960s C.L. Voorhoeve did extensive survey work on languages of the Trans Fly region and the south-west lowlands of Irian Jaya, and was also able to draw on a pioneering comparative study by Healey (1964) on the Ok family of central New Guinea. Voorhoeve found resemblances in lexicon and grammar between five small language families and several isolates which led him to posit the Central and South New Guinea (CSNG) stock, comprising about 60 languages in all (Voorhoeve 1968). During the same period Kenneth McElhanon (then a PhD student at ANU) and associates of the Papua New Guinea branch of the Summer Institute of Linguistics surveyed languages of Morobe Province and the eastern part of Madang Province. Using this data McElhanon (1967, 1970, 1975) proposed a genetic group consisting of some 70 languages, which he called the Finisterre-Huon phylum. When the two scholars compared notes in the late 1960s they found about 90 possible cognates in basic vocabulary common to the CSNG and Finisterre-Huon groups.<sup>4</sup> They posited a remote genetic relationship between the two and named the wider group the Trans New Guinea phylum (McElhanon & Voorhoeve 1970). I will refer to their proposed grouping as *Trans New Guinea I*.

---

<sup>4</sup> McElhanon and Voorhoeve’s cognate sets represent only 53 meanings from the basic vocabulary list. However, they found two separate series of cognate sets for many meanings, yielding some 90 putative cognate sets in all.

The fact that Finisterre-Huon languages are spoken in north-east New Guinea, mainly in Morobe Province, and CSNG languages are spoken hundreds of kilometres to the west, and that the two are separated by many intervening linguistic groups, strongly suggested that Trans New Guinea I might prove to be the nucleus of a larger genetic stock. In fact, Wurm had earlier posited a genetic relationship between his East New Guinea Highland stock and a handful of languages lying to the immediate south and south-west of the central highlands and had voiced suspicions of a relationship between the East New Guinea Highlands stock and the Huon Peninsula and Ok groups (Wurm 1964, 1971). McElhanon and Voorhoeve (1970) foreshadowed certain extensions, noting that some Finisterre-Huon and CSNG cognate sets have resemblant forms in other Papuan languages. In their map they included in the TNG phylum the Binandere group of south-east Papua (see Map 1) and indicated that other groups such as the East New Guinea Highlands stock posited by Wurm (1964, 1965), the Madang phylum of eastern Madang Province proposed by Z'graggen (1971) and the small Nimboran group of north-east Irian Jaya, would probably turn out to belong as well.



**Map 1:** Trans New Guinea I (after McElhanon & Voorhoeve 1970)

The terms ‘stock’ and ‘phylum’ were used in a specialised sense by McElhanon and Voorhoeve. They, like Wurm, adopted a system of nomenclature that ranks the closeness of languages in terms of gross number of lexical resemblances (by percentage of putative cognates in basic vocabulary). In this system a phylum is a collection of languages whose most distantly related members share between 5 and 12 per cent of cognates on the 200-item list. Lower order groups are ‘stocks’ (about 12–28 per cent cognation), ‘families’ (28–45), ‘subfamilies’ (45–81) and ‘languages’ (consisting of dialects sharing more than about 81 per cent of cognates). Those, at any rate, were the ideal ranges. In practice, the percentages allowed for acceptance into a phylum were often less than five per cent. And a

new level, that of ‘macro-phylum’ was added so that two or more phyla could be connected. However, the ‘macro’ was not added consistently.

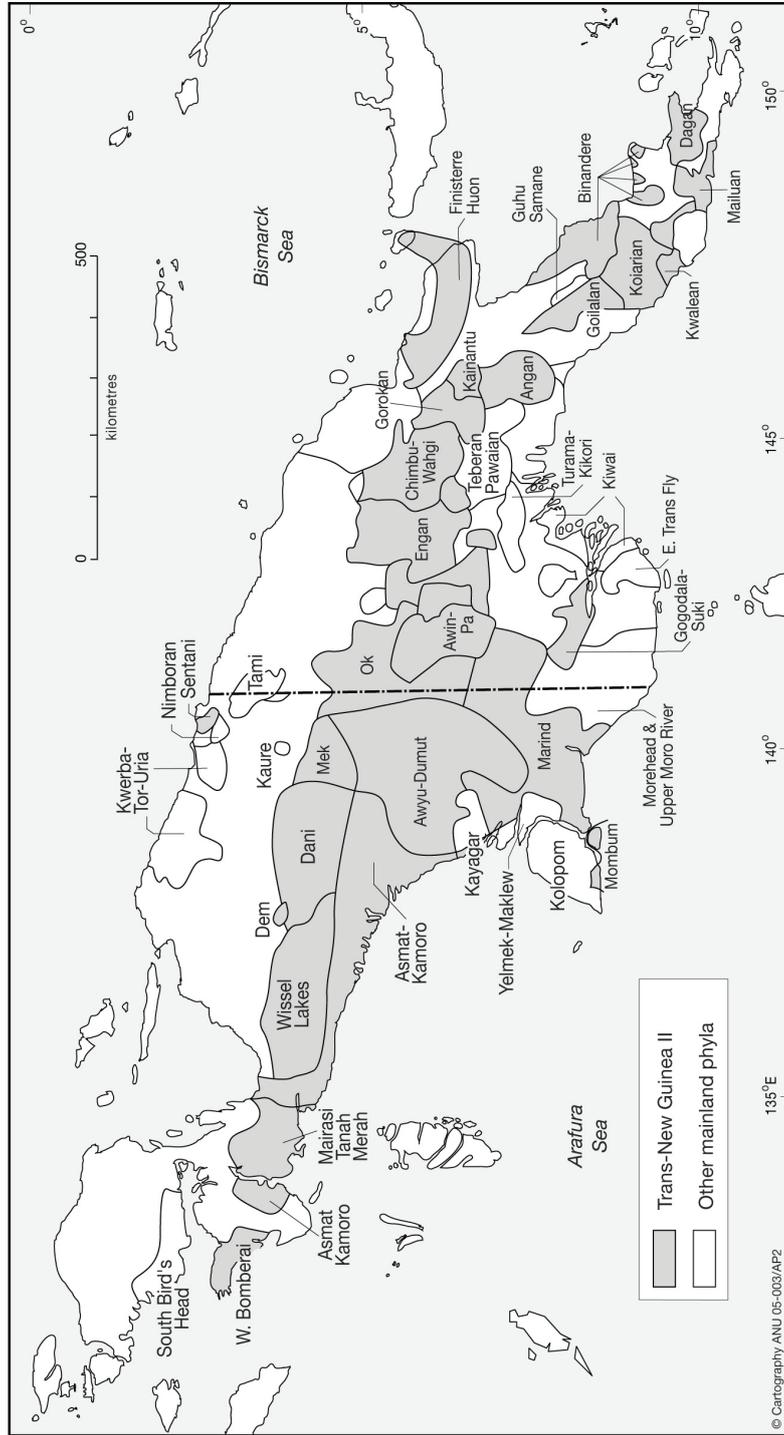
Although this nomenclature has the virtue of giving a rough indication of degrees of lexical diversity it is not used by most comparative linguists. The problem is that lexicostatistics, as normally applied, is a very blunt instrument and often yields classifications which are at odds with those based on the shared innovations method. The standard practice is to use ‘family’ for the largest well-established genetic group, with its subgroups simply termed ‘first order’, ‘second order’, and so on. In this paper I follow standard practice, for example I will speak of the Trans New Guinea ‘family’, except when quoting works that adopt the lexicostatistical nomenclature.

### **2.3 Trans New Guinea II**

Between 1970 and 1975 the ANU research group greatly extended the membership of the proposed TNG family. As diagnostics for deciding whether a language belongs in TNG, Wurm and McElhanon (1975:150–151), Wurm, Voorhoeve and McElhanon (1975:306–307) and Wurm (1975a, b) referred to four main types of evidence:

- (a) The matrix of lexicostatistical comparisons.
- (b) A small number of putative cognate sets (about 10) in basic lexicon which have wide distributions and seem to be very stable. If a language has several forms belonging to these cognate sets it is assigned to TNG.
- (c) Three sets of free form pronouns, called 1, 2 and 3, were posited as having great antiquity in Papuan languages. Reflexes of set 1 were said to be confined to TNG. If a language has several pronouns belonging to set 1, especially the 1st, 2nd and 3rd singular and 1st plural forms, it is assigned to PNG. Reflexes of sets 2 and 3 are found both inside and outside of TNG and are not considered diagnostic of TNG membership. In fact, if a language has some set 2 or set 3 forms, these are taken as evidence of a non-TNG substratum.
- (d) Certain striking structural features in morphology and syntax are common among TNG languages and rare in other Papuan languages. Possession of these structural features was regarded as an indicator of membership in TNG, although not such a strong one as (b) or (c).

At least two different kinds of extended TNG groupings were posited in the various papers in Wurm (ed. 1975). The nature of the difference is made explicit in the central paper by Wurm, Voorhoeve and McElhanon. They referred to a ‘main section’ or core group made up of about 256 languages, whose membership was regarded as quite secure. Let us call this group *Trans New Guinea II*. It contains all the languages of the central cordillera east of the Bird’s Head, from the Wissel Lakes and the Baliem Valley to south-east Papua, together with some languages spoken to the north of the central ranges (chiefly the Finisterre-Huon and Binandere groups) and a few spoken to the south (chiefly the Asmat-Kamoro, Awyu-Dumut and Lowland Ok groups).



**Map 2: Trans New Guinea II (after Wurm, Voorhoeve & McElhanon 1975)**

## 2.4 Trans New Guinea III

Wurm, Voorhoeve and McElhanon (1975) also posited a much larger, more speculative group, which I will refer to as *Trans New Guinea III*. This contains some 491 languages. It comprises TNG II plus another 235 or so languages whose claims were considered by Wurm, Voorhoeve and McElhanon (1975) to be marginal, including the members of the following groups: Madang-Adelbert Range, Border, Eleman, Inland Gulf, Kalam, Kolopom, South Bird's Head, South-east Papuan, Teberan-Pawaian and Trans Fly (see Map 3). The Papuan languages of Timor, Alor and Pantar in eastern Indonesia were also included in TNG for the first time, but in the marginal category. All these languages were considered to be marginal because they appeared in certain respects to be structurally aberrant—exhibiting 'non-TNG phylum' features—and because (in many cases) they also seemed to show very few specific lexical resemblances with typical TNG languages. Wurm, Voorhoeve and McElhanon (1975) regarded the structurally aberrant languages as hybrids, resulting from the overlaying of a TNG component on a non-TNG substratum (see under (5) in section 2.5).

TNG III covers most of the inhabited regions of the New Guinea mainland. Almost the only Papuan languages of New Guinea excluded from it were (i) most of those spoken in the Bird's Head and Bomberai Peninsula at the western end of the island and (ii) many of the languages of those parts of central New Guinea that lie to the north of the central cordillera, especially in Sandaun and East Sepik provinces and in the western part of Madang Province.

Two influential reference works that appeared during the next few years—the two-volume *Atlas of languages of the Pacific* (Wurm & Hattori 1981–1983) and *The Papuan languages of Oceania* (Wurm 1982)—made the findings of the 1975 book accessible to a wider public but did not add significantly to the evidence. A bit surprisingly, it is the larger and more speculative group, TNG III, rather than the more conservative TNG II, that is represented in the main maps of these two works, as well as in Wurm (ed. 1975). Because of these works, and especially the Atlas, the major genetic groupings of Papuan languages proposed in Wurm (ed. 1975) gained a considerable measure of acceptance by non-specialists.

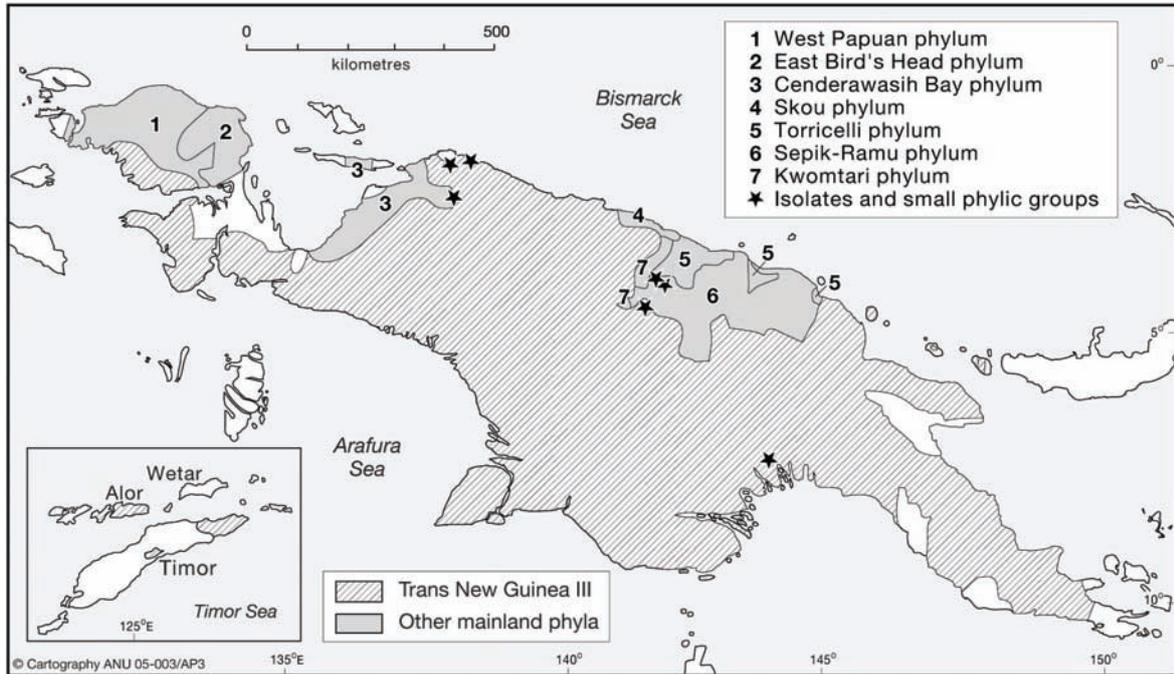
## 2.5 A critique of the case for TNG I–III

The case for a Trans New Guinea family was not widely accepted by linguists working in Papuan comparative studies. While praising the enormous amount of pioneering comparative work accomplished by the ANU team, all the main reviewers of Wurm (ed. 1975) regarded the TNG hypothesis as unproven, although by no means without promise (Foley 1986; Haiman 1979; Heeschen 1978; Lang 1976). Indeed, serious reservations were entered by two of the principal contributors to Wurm (ed. 1975), namely McElhanon (1975) and Z'graggen (1975).<sup>5</sup>

---

<sup>5</sup> McElhanon writes of the tentative nature of many of the subgroups, adding that the use of terms such as 'phylum' and 'stock'

does not imply that the present writer regards such groups as having been established. Rather he takes the view that the [TNG] phylum has been posited and that the phylum will be confirmed and its subgrouping established only after a rigorous application of the comparative method (McElhanon 1975:528).



**Map 3:** Trans New Guinea III (after Wurm & Hattori 1981–1983)

While these critics were principally addressing the extended versions of the hypothesis—TNG II and TNG III—it is clear that some also had misgivings about TNG I. Foley (1986) was not prepared to accept that a convincing case had been made for TNG or indeed for any family of Papuan languages larger than about 30–40 languages. The objections to TNG I–III fall into several categories.

### (1) The comparative method was not properly applied

Over a century of work on such families as Indo-European and Austronesian has shown that by far the most reliable method for demonstrating genetic relationships between languages, for determining the subgrouping or genealogical structure of a family, and for reconstructing the historical phonology and lexicon of earlier stages is the so-called ‘comparative method’, supported by the methods of dialect geography.

At one level the ‘comparative method’ is properly viewed as a theory about how certain kinds of continuity and change in languages occur. At another level it is a method, a set of procedures for organising and evaluating data in the light of the theory. The fact of nature which makes the comparative method so powerful is this: sound changes, such as  $p > f$ ,  $t > s$  before  $i$ , or loss of  $h$ , typically are unconsciously made in the speech of individual native speakers, and by and large are regular across a definable speech community. The method

---

Z’graggen questions the adequacy of the crucial comparisons made by McElhanon and Voorhoeve (1970). He points to the many unexplained elements attributed to old compounds and to other unexplained irregularities in the putative cognate sets. He notes that of the 93 sets of forms, only 10 per cent are found in more than half the languages compared and he concludes that ‘In the writer’s opinion, the evidence listed may be inadequate for the postulation and illustration of a genetic relationship in the traditional way’ (Z’graggen 1975:587).

requires us to search in two or more languages for regular sound correspondences in arbitrary (that is non-onomatopoeic) form-meaning pairings including basic vocabulary and paradigmatic sets of grammatical elements. Such correspondences provide a principled basis for reconstructing the sound system, lexicon and morphology of the common ancestor. These reconstructions in turn make possible a search for innovations, on which a reliable subgrouping must rest, and for distinguishing borrowed from inherited material. To apply the comparative method thoroughly takes a long time and good descriptive data are needed.

Was the comparative method properly applied in Wurm (ed. 1975)? The consensus is that it was not, in those crucial papers that deal with reconstructions at the TNG level and in tests to determine whether particular languages are TNG or not. For example, in papers by Wurm (1975a), Wurm and McElhanon (1975) and Wurm, Voorhoeve and McElhanon (1975) it is noted that a handful of etyma are fairly stable in TNG languages. But nowhere is a systematic attempt made to work out the phonological history of the putative cognate forms, by establishing regular sound correspondences that would underpin the reconstruction of a phonological system for Proto Trans New Guinea (pTNG), and so permit specific lexical form to be reconstructed. This lack led to a bizarre situation whereby the authors used English glosses ('eat', 'say/speak', 'louse', 'bone', 'head', and so on) rather than pTNG forms to refer to putative pTNG lexemes. The lack also meant that the authors had no clear grounds for distinguishing between common inheritance, borrowings and chance resemblances as explanations for resemblant forms.

The critics complained about the fact that, insofar as any reconstruction was attempted in the book, it was largely 'top-down'. That is to say, distantly related languages were compared to arrive at rough approximations of pTNG forms, without first preparing the ground by reconstructing from the bottom up, that is beginning with lower order groups. Lang (1976:74) described as 'near impossible' the task of establishing a pTNG sound system 'without the reconstruction of intermediate systems'. Haiman (1979:897) concluded that 'the reconstruction of [Trans New Guinea Phylum] forms is not a useful task at this time', meaning top-down reconstruction would not yield reliable results. It seems that both Haiman and Lang favoured an exclusive concentration on lower-level reconstruction in the first instance.

## **(2) Undue weight given to typological resemblances**

The critics considered that the proponents of TNG had placed too much emphasis on structural resemblances as diagnostics. It is generally agreed in historical linguistic circles that many kinds of structural features can be readily borrowed and some kinds are likely to develop independently. In general, then, structural resemblances do not constitute strong evidence for genetic relationship unless they are also associated with cognate morphemes. None of the structural features has a geographic distribution that coincides exactly with that of TNG II or TNG III. Perhaps the most striking structural feature associated with TNG II and III is switch reference marking, with 'medial' verbs marked for same or different subject as the next verb and (often) for relative tense. The various sorts of switch reference systems found in New Guinea, and their geographic distribution, are described in Roberts (1988, 1997). It turns out that while there is a reasonably close match, not all TNG II and III languages exhibit switch reference marking while such marking is found among some languages that are excluded from these groups. I will return to this matter in section 3.5 below.

### **(3) The lexicostatistical evidence is unconvincing**

The lexicostatistical agreements between distant branches of TNG II and TNG III are very low. Wurm (1971:585) says that the languages in the East New Guinea Highlands stock, the Huon stock, the Central and South New Guinea stock, and the West New Guinea Highlands phylum show an average of 3–7 per cent cognation, while these groups and the South-east Papuan phylum show an average of 2–3 per cent. Three objections can be raised against such low agreements as grounds for claiming genetic relations: (1) We are not dealing with established cognates here, but with ‘resemblant forms’, which are *possibly* cognate. (2) Chance resemblances can produce percentages of resemblant forms common to unrelated languages of at least four per cent. (3) Some of the resemblant forms may be borrowings.

Objections (1) and (3) also apply to lexicostatistically-defined subgroups of TNG whose members show much higher agreements. The matter of borrowing will be raised again under point (5) below.

### **(4) Failure to exploit morphological paradigms**

The ANU group were in fact holding an ace but they did not play it effectively. The strongest grounds they had for a large TNG group lay in the pronominal evidence. Voorhoeve (1975) and Wurm (1975b) noted that widespread cognates in certain independent pronoun forms allow the reconstruction of a near complete paradigm. But although approximate reconstructions of five or six forms in this paradigm were made, the supporting evidence was not assembled and used to trace the historical development of pronouns across the TNG phylum. Instead Wurm’s (1975b) discussion of pronouns muddied the waters with a liberal infusion of substrate theory (Voorhoeve 1987; Ross 1995). This account puts forward several questionable ideas: that TNG languages reflect one or another three originally distinct pronoun sets, or some combination of these, that all three are ancient Papuan sets, originally belonging to unrelated genetic stocks and that their occurrence of sets 2 and 3 in TNG languages is evidence that these languages were hybrids, with a TNG overlay on a non-TNG substratum. Regardless of the truth or otherwise of these proposals, the case for TNG would have been significantly strengthened by basing reconstructions on a systematic comparison of contemporary data.

### **(5) Is the family tree model appropriate? The question of genetic continuity in TNG III languages**

The important role that Wurm and his co-authors attribute to substrata substantially weakens their case for an extended TNG genetic group. It raises the question whether the genetic (or family tree) model is appropriate for making sense of the historical relationships of Papuan languages. Some of the key papers in Wurm (ed. 1975) took an equivocal stance on this point.

The family tree model holds that languages are stable codes typically learnt by successive generations of native speakers. We can speak of genetic continuity so long as the line of native speaker transmission is unbroken. Linguistic splitting occurs when a population of native speakers that once spoke the same language become separated by geographic or social barriers and the two daughter communities evolve their codes

independently. A family tree diagram schematically represents a sequence of such purported splits.

The following statement indicates that Wurm et al. regarded some languages—those that are typologically and lexically most divergent from the rest—as being members of the TNG family only in a secondary sense, as a result of non-native speakers adopting a pidgin form of TNG in place of their mother tongue.

[I]t appears that much of the Trans New Guinea Phylum area may have originally been occupied by a number of probably unrelated earlier languages, and that *the inter-relationship of many of the present-day Trans-New Guinea Phylum languages is, in a way, secondary, or partial and fractional, in nature* and brought about by the very strong and pervading influence of an originally little differentiated element manifested in both the lexical and structural-typological levels, and attributable to the spreading of daughter languages of the Trans-New Guinea Phylum proto-language first from west to east through much of the New Guinea mainland well over five thousand years ago, and perhaps much more vigorously, from east to west during the last five thousand years or so ... *The presence of the older, different languages upon which the Trans-New Guinea Phylum languages appear to have been superimposed in the course of these migrations, is noticeable in the form of substrata of varying strength throughout the greater part of the Trans-New Guinea Phylum.* (Wurm, Voorhoeve & McElhanon 1975:300; **my italics:AP**)

To explain the distribution of linguistic features in terms of a substrate (pre-TNG) language having a superstrate (TNG) imposed on it is, as I understand it, to propose a sequence of developments of the following sort. First, native speakers of language A (pre-TNG) come in contact with an invasive or neighbouring language, B (TNG). Speakers of A learn an imperfect version of B, with much interference from their native tongue in pronunciation, grammar, and so on. Let us refer to any version of language B imperfectly learnt by speakers of A as ‘Bx’. Bx is a suddenly-formed hybrid language, which is not the direct descendant of any one language. The community whose mother tongue is A at first uses Bx only as a lingua franca or contact language with their neighbours. Bx has no native speakers and is in fact not a stable language but more of a collection of idiolects which vary markedly from one speaker to another. Later, however, the speakers of A (or their children or children’s children) give up A completely in favour of Bx. When this happens, Bx becomes a stable language which is however full of substrate features derived from A. This kind of language shift, implied by Wurm et al., has very little support in the literature.

Wurm et al. go on to indicate the regions of New Guinea where substrate residues are strongest:

The main characteristics [of TNG phylum languages] show a fair amount of homogeneity ... except that the influence of various substrata is in evidence in most parts of the phylum, with their influence being particularly strong in some, mostly marginal areas where the languages contain a considerable number of non-Trans-New Guinea Phylum features and are quite aberrant ... Such areas are, in particular, in a rather extended region in the central south, in the border area between the West Sepik District [today’s Sandaun Province] of Papua New Guinea and Irian Jaya, in the north and extreme west of Irian Jaya, as well as in Madang District ... *it has nevertheless been decided to include such fringe area language groups of the Trans-New Guinea Phylum, ... even though only a component part of each of them is likely to be genetically related to other Trans-New Guinea Phylum languages.* Other language groups which ... show quite strong, but apparently less incisive, Trans-New Guinea

Phylum influence, have been excluded ... with the decisions ... being perhaps somewhat arbitrary in some cases. (Wurm, Voorhoeve & McElhanon 1975:300; **my italics:AP**)

It seems that language shifts of this kind are rare and have been observed to occur only in rather special social circumstances. Those well-known creole languages whose origins are associated with plantation societies represent a special type of hybridisation, one in which speakers of several or many substrate languages in a heterogeneous migrant community are simultaneously exposed to a prestige language which they must learn for their survival.

In his review of Wurm (ed. 1975), Ranier Lang was sharply critical of the weight given to substratum influence as an explanation of diversity within TNG languages. He argued that:

what evidence we have of population movements in Papua New Guinea is of a kind that does not allow for substrata. Populations have been displaced in recent history ... through either of two events (or a combination of the two): (a) natural disasters such as volcanic eruption, an earthquake, or drought and/or frost have driven populations from their home ground; (b) warfare has had the same effect. When they have left their home ground they have either moved into virgin bush to carve out an entirely new existence for themselves ... or they have taken refuge with allies, in which case they have been absorbed into the host group, thus giving up their language and adopting that of their hosts.

... It would seem that the natural fragmentation of the country and the social conditions (partly brought about by geographical factors) would be much stronger determinants of linguistic diversity than substratum influence. But how the social conditions bring about linguistic changes, of this we know precious little in the New Guinea area. The sophisticated sociolinguistic research has just not yet been carried out. (Lang 1976:77–78)

#### **(6) Has accumulated borrowing made it impossible to determine deep genetic relationships among Papuan languages of New Guinea?**

Whatever one thinks of the substratum argument it must be conceded that borrowing has played a rather important role in shaping the lexicon of many New Guinea languages. While ‘basic’ vocabulary, in general, is less prone to borrowing than ‘cultural’ vocabulary the difference is only one of degree.

I will mention just two studies indicating extensive borrowing of basic vocabulary between neighbouring Papuan languages that are only distantly related. Comrie (1986, 1989) found that Haruai, an unplaced language of the Western Schraders, shares about 35 per cent resemblant forms with Kobon, a neighbouring language belonging to the Kalam branch of the Madang group. Given that the genetic relationship between Haruai and Kobon is extremely remote (they are very unlike in morphology) almost all of this agreement can be attributed to borrowing. In a similar vein, Shaw (1986) notes that Huli, a language of the Engan group spoken in the Southern Highlands of PNG shares some five to ten per cent of resemblant forms on a basic vocabulary list with Bogaya, a language of the Central and South New Guinea stock spoken not far away in the Mount Bosavi region. Duna, another language of the Southern Highlands Province, was found by Shaw to share 27–32 per cent resemblant forms with Huli and 20–28 per cent with Bogaya. It would seem that Duna’s percentages with either Huli or Bogaya (or with both) have been inflated by about 20 per cent by borrowing.

In *The Papuan languages of New Guinea* William Foley questions the feasibility of applying the comparative method to the lexicon of Papuan languages, except when the languages are quite closely related, say, as the languages of the Germanic family are, or those of the Romance family—with a common ancestor spoken no more than about 2000 years ago (Foley 1986:213, 228). Foley suggests that there are demographic and social factors that make Papuan languages particularly difficult subjects for the comparative method at greater time depths.

Papuan language families are small and are generally spoken in small areas. The languages are usually contiguous, and have been so for millennia. None of the particular historical and geographical patterns necessary for the smooth application of the comparative method obtain in Papuan languages. Rather, ... Papuan languages normally exhibit a pattern of enormous cross-influence in all areas; so in no sense can the assumption that the daughter languages develop independently be taken as viable in this context. As the comparative method, with its sorting of cognates from borrowing, is deeply grounded in the family tree model, its application to Papuan languages is no mean problem, and suggests that some major rethinking of the method itself may be needed for these languages. (Foley 1986:209–210)

At times in his discussion Foley seems to take a position similar to that advanced by Stephen Wurm in a number of works and more recently by R.M.W. Dixon when writing about comparative linguistics and the Australian languages (Dixon 1997, 2002). Foley asks:

Given the extensive borrowing that has been demonstrated ... the question of the meaning of genetic affiliation in these languages presents itself. Could borrowing proceed on such a scale that a language could be said to be mixed? (Foley 1986:213)

Here the point needs to be made that the ‘mixing’ of languages by the cumulative effects of diffusion is a very different thing from genetic discontinuity, in which a generation of speakers suddenly creates a new language by ‘mixing’ elements of two or more existing languages. Diffusion occurs when speakers of language A borrow or copy word forms and syntactic or semantic patterns from language B, while both languages continue to be spoken as stable codes. There is no break in native-speaker transmission. Speakers of A do not give up their mother tongue following contact with B. They continue to speak it, with the addition of some borrowed elements. Even though English has absorbed extensive loans from Scandinavian, French and other languages along the way, Modern English is descended in an unbroken line of transmission from the Anglo-Saxon dialects which entered Britain in the 5th and 6th centuries AD.

Foley objects to this position in the following terms:

Of what use is it to say that language X belongs to family Y because proto-Y was its ancestral starting point, if most of its vocabulary and structure actually come from family Z? To ignore all the features of Z found in X, just to get a simple yes/no classification of X in family Y, ignores all the fundamental historical processes that have been at work in Y. (Foley 1986:213)

However, this objection targets a straw man. A linguist applying the comparative method cannot properly ignore borrowings en route to arriving at a genetic classification or a subgrouping. One of the first tasks is to examine a representative sample of the whole vocabulary of each language, and to look for ways and means of distinguishing between directly inherited elements and borrowed elements. In the trickier cases a reliable genetic classification cannot be achieved until that sorting out is done. To put it another way, the

task establishing the main line of genetic continuity is a central part of unravelling ‘the fundamental historical processes that have been at work in Y’. One cannot get far in such an enquiry without first being able to distinguish layers of borrowing from the directly inherited component. A genetic classification of Y is thus very useful, even if it is not the whole story.

A closer reading of Foley suggests that he was not denying that contemporary Papuan languages do not have long genealogies. He was merely pessimistic about the chances of being able to prove lines of genetic continuity going back well beyond 2000 years, by the standard method of using cognate lexical items to reconstruct the phonology and morphology of earlier stages. Instead, Foley pins his hopes on shared morphology rather than shared lexicon as the best evidence for establishing distant genetic relationships. In the case of TNG I think it can be shown that his pessimism concerning the feasibility of using lexical cognates to establish deep relationships, while understandable, was excessive.

### **Summing up**

Were the critics right to reject the TNG hypothesis (in any of versions I–III) as unproven? I believe they were. That is not to say that the total body of evidence was weak. I believe the problems lay more in the way the evidence was selected and presented in the 1975 book and associated works, and in the ANU researchers’ failure to proceed logically to the next phase of analysis. Having put a lot of energy into surveys and into quick searches for clues as to genetic connections, the research group did not carry the job through by systematically applying the comparative method. In fact, they showed a mysterious reluctance to do so.

## **3 A current view of the Trans New Guinea hypothesis**

### **3.1 Trans New Guinea IV**

Recently the TNG hypothesis has been put under scrutiny again, using classical comparative methods. Over the last few years Malcolm Ross, Meredith Osmond and I have begun to sift through the growing body of descriptive data on Papuan languages, using both top-down and bottom-up approaches. Our project owes much to the pioneering comparative studies of the 1960s and 1970s, not in the matter of method but in their pointers to relevant evidence. At first we concentrated on the 100-member subgroup recognised by Z’graggen (1975, 1980a-d) which he named the ‘Madang-Adelbert Range sub-phylum’ (we refer to this simply as the ‘Madang’ group). Later we extended the scope of the study to all putative TNG groups.

This recent work supports a modified form of the TNG hypothesis. At the risk of confusing the reader, who already has had to contend with three variants, we will call our version *Trans New Guinea IV*. There would be a greater risk of confusion if the modified grouping were left unlabelled. The boundaries and subgroups of TNG IV are shown in Map 5 in Malcolm Ross’s chapter in this volume. The membership of this revised grouping will be discussed further in section 4.1; see also Ross’s chapter. Suffice to say here that TNG IV includes the large Madang-Adelbert Range group, the Papuan languages

of Timor, Alor and Pantar and the small West Bomberai group, and *most* of the lower order groups assigned to TNG III.

The main evidence for TNG IV consists of:

- (i) A body of regular sound correspondences, based on (ii), which has allowed a good part of the Proto TNG sound system and its development in a sample of eight daughter languages to be reconstructed. See brief discussion in section 3.2.
- (ii) About 200 putative cognate sets represented in two or more major subgroups. Most of these sets refer to near universal concepts (so-called ‘basic vocabulary’). pTNG reconstructions based on these are listed in section 3.3.
- (iii) Systematic form-meaning correspondences in the personal pronouns, permitting reconstruction of virtually a complete paradigm (see section 3.4).
- (iv) Widespread resemblances in fragments of certain other grammatical paradigms (see section 3.5).

A small sample of TNG cognate sets is presented in Table 1.

**Table 1:** Some cognate sets of the Trans New Guinea family

|                                | ‘breast’ | ‘eat’ | ‘louse’ | ‘name’ |
|--------------------------------|----------|-------|---------|--------|
| pTNG                           | *amu     | *na-  | *niman  | *ibi   |
| Asmat (Irian Jaya)             |          | na-   |         | yipi   |
| Kiwai (SW coast, PNG)          | amo      |       | nimo    |        |
| Kewa (W. Highlands, PNG)       |          | na-   |         | ibi    |
| Kuman (C. Highlands, PNG)      | aemu     |       | numan   |        |
| Kube (Morobe, PNG)             | namu     | ne-   | imiŋ    |        |
| Katiati (Madang Province, PNG) | ama      |       | ñima    | nimbi  |
| Aomie (Central Province, PNG)  | ame      |       | ume     | ihe    |

In addition, the distribution of certain striking structural resemblances noted by Wurm and others has been more precisely charted, and shown to correlate rather closely with the distribution of TNG IV languages. While such structural evidence cannot be primary grounds for positing a genetic stock it carries some weight as corroborative evidence. The following sections will deal briefly with each of these categories of evidence.

In one respect our experience contradicts the views expressed by several of the critics of Wurm (ed. 1975) that, in the TNG case, top-down reconstruction is impractical or pointless until a solid base of bottom-up reconstructions have been obtained. It is true that there are many apparent irregular correspondences that can only be made sense of by bottom-up comparisons in conjunction with careful study of dialect geography. But substantial initial progress can be made by top-down comparison, if this is done methodically (Pawley 2001; Ross 2000). In fact, bottom-up comparisons have proved difficult to do well without top-down reconstructions available as a check. It is best to work in both directions at once.

### 3.2 A tentative reconstruction of Proto TNG segmental phonology

Space does not allow anything more than a fragmentary exemplification of the evidence for pTNG phonology, some of which is presented in Pawley (1995, 2000a, 2001, in prep.).

The following consonant and vowel phonemes can be reconstructed for proto Trans New Guinea (pTNG), based on correspondences in a set of cognates between languages drawn from several diverse subgroups.

**Table 2:** pTNG segmental phonemes

| <b>consonants</b>       |          |         |         |       |
|-------------------------|----------|---------|---------|-------|
|                         | bilabial | apical  | palatal | velar |
| oral obstruents         | p        | t s     |         | k     |
| prenasalised obstruents | mb       | nd      | ɲj      | ŋg    |
| nasals                  | m        | n       |         | ŋ     |
| lateral                 |          | l       |         |       |
| glide                   | w        |         | y       |       |
| <b>vowels</b>           |          |         |         |       |
|                         | front    | central | back    |       |
| high                    | i        |         | u       |       |
| mid                     | e        |         | o       |       |
| low                     |          | a       |         |       |

The reconstructed consonants and vowels in Table 2 are not an exhaustive list of the sets of distinct or partially distinct correspondences represented in our data. These symbols simply represent a list of best-attested correspondence sets, which yield a plausible phonological system. There remains a large residue of more problematic correspondence sets to be dealt with, as well as the question of whether pTNG had tonal contrasts, as a number of its daughter languages do. Many of these problematic correspondences should be amenable to explanation in terms of natural processes of sporadic change (assimilation, dissimilation, and so on). However, there is some evidence for attributing to pTNG a second lateral phoneme \*L, and a contrast between \*t and \*r. \*L is used below in certain reconstructions where some correspondences point to \*t and others to \*l.

The pTNG nasals \*m and \*n are well attested, being regularly reflected in dozens of cognate sets in both word-initial and -medial position. These are the two most stable phonemes in terms both of continuity of contrasts with other phonemes and continuity of phonetic character. The velar nasal \*ŋ looks fairly secure even though there are only three or four good cognate sets supporting it. Some TNG languages show a fourth nasal consonant, palatal [ɲ], and a prenasalised palatal affricate, [ɲj], although in at least some cases these may be post-pTNG developments.

It is clear that in pTNG there was a contrast between obstruents in the bilabial, apical and velar positions. It is also clear that there were two contrasting series of obstruents and it is very likely that the basic opposition was between prenasalised and oral obstruents.

The symbols for particular proto-phonemes should not be taken to have constant phonetic values. Many TNG languages show fairly considerable allophonic variation in obstruent phonemes. A phoneme written /b/ may have variants [b, mb, mp], one written /g/

may have variants [g, ŋg, ŋk], one written /p/ may have variants [p, f, v], and so on. Many languages have no contrast between [t] and flapped [r], some do not distinguish between [t] and [s], or between [y] and [dz].

The three vowels \*a, \*i, \*u are very well attested, \*e and \*o less so. The symbol \*V is used to represent an indeterminate vowel, that is a set of vowel correspondences not readily assignable to any of the five attested sets. It is possible that there were additional vowels but no very clear patterns of correspondences have yet emerged among that residue of material that does not fit the five vowel correspondences.

### Syllable and word base structure

Syllables had the shape (C)V, and in word-final position they could also be (C)VC. There were probably no phonemic consonant clusters within words, phonetic homorganic nasal + obstruent clusters being interpreted as unit phonemes. Lexical bases (morphemes) could consist of one or more syllables, for example \*na- ‘eat’, \*nVŋg- ‘know’, \*imbi ‘name’, \*niman ‘louse’, \*takVn[V] ‘moon’, \*mangat[a] ‘teeth’, \*kumbutu ‘wind’, \*tutu(tu)ga ‘straight’. Many verb stems are monosyllabic. The status of stress and tone remains uncertain.

The following are reflexes of the pTNG consonants and vowels reconstructed in Table 2 in one daughter language: Kalam, of the Schrader Ranges, Madang Province:

### Obstruents

- \*mb > b (realised as [mb] initially and medially, [mp] finally): \*mbapa ‘father’ > bapi, \*ambi ‘man’ > b, \*sambV ‘cloud’ > seb, \*imbi ‘name’ > yb, \*kamb(a,u)u[na] ‘stone’ > kab, \*si(m,mb)i ‘guts’ > sb
- \*mb > m in at least one case: \*mbalaŋ ‘flame’ > malaŋ, maŋlaŋ. Note also \*(mb,m)elak ‘light, lightning’ > melk ‘light’
- \*p > p initially and medially (realised as [ϕ] initially, [β] medially): \*panV ‘female’ > pañ, \*apus(i) ‘grandparent’ > aps ‘grandmother’, \*[ma]pVn ‘liver’ > mapn
- \*nd > d [nd] medially: \*mund-maŋV ‘heart’ > mudmagi, \*kindil ‘root’ > kdl
- \*t > t initially and finally (realised as [t] initially, [r] elsewhere): \*takVn[V] ‘moon’ > takn, tuk- ‘cut’ > tk- ‘sever’, \*tu ‘axe’ > tu, \*tumuk or kumut ‘thunder’ > tumuk, \*-iL ‘2/3 dual verbal suffix’ > -it
- \*t > zero medially or finally in one case: \*mangat[a] ‘teeth’ > meg
- \*s > s initially and medially: \*sambV ‘cloud’ > seb, \*si(m,mb)i ‘guts’ > sb, \*[si]si ‘urine’ > ss; \*apus(i) ‘grandparent’ > \*aps ‘grandmother’
- \*ŋg > g [ŋg] medially and [ŋk] finally: \*mangat[a] ‘teeth’ > meg, \*maŋV ‘round object, egg’ > magi. In one case \*g has varying reflexes in different dialects of Kalam: \*nVŋg- ‘see’ > ng-, nŋ- in Ti dialect, but nŋ- in Etp dialect.
- \*k > k ([ʏ] medially, [k] elsewhere): \*kambu[na] ‘stone’ > kab, \*ka(w)nan ‘shadow’ > kawnan, \*kinV- ‘sleep’ > kn-, \*kumV- ‘die’ > kum-, \*kakV- ‘carry on shoulder’ > kak-, \*muk ‘milk’ > muk (Ti dial. mok), \*muk ‘brain’ > muk, \*takVn[V] ‘moon’ > takn, \*tuk- ‘cut’ > tk- ‘sever’

**Nasals**

- \**m* > \**m* : \**mangV* ‘egg’ > *magi*, \**ma-* ‘not’ > *ma-*, \**m[(o,u)k* ‘milk’ > *muk, mk*, \**mVna-* ‘be’ > *md-*, \**am(a,i)* ‘mother’ > *ami*, \**kuma-* ‘die’ > *kum-*, \**niman* ‘louse’ > *yman*
- \**n* > *n* : \**nVng-* ‘see’ > *ng-*, *nŋ-*, \**kinV-* ‘sleep’ > *kn-*, \**niman* ‘louse’ > *iman*, \**takVn[V]* ‘moon’ > *takn*, \**wani* ‘who?’ > *an*
- \**n* > *ñ* in a few words: \**nV* ‘child’ > *ñi* ‘son’, \**panV* ‘female’ > *pañ* ‘daughter’, \**nok* ‘water’ > *ñg*; in one case \**n* > *d* : \**mVna-* ‘be’ > *md-*
- \**ŋ* > *ŋ*. Attested only in medial and final position: \**nVng-* ‘see’ > *ng, nŋ-*, \**ŋaŋa* ‘baby’ > *-ŋaŋ*, \**balan* ‘flame’ > *malan, manlan*

**Other resonants**

- \**l* > *l* (retroflex lateral): \**balan* ‘flame’ > *malan, manlan*, \**[mb, m]elak* ‘light, brightness’ > *melk*, \**kindil* ‘root’ > *kdl*
- \**w* > *w* although good cognate sets are few: \**k(aw,o)nan* ‘shadow/spirit’ > *kawnan*, \**walaka* ‘testicles’ > *walak*, \**wati* ‘fence’ > *wati*
- \**y* > *y* but the relevant cognate sets are few: \**yaka[1]* > *yakt* ‘bird’, \**aya* ‘sister’ > *ay*

**Vowels in stressed position**

The most common Kalam outcomes of pTNG vowels in stressed position are as follows:

- \**a* usually gives *a* : \**am(a,i)* ‘mother’ > *ami*, \**kakV-* ‘carry’ > *kak-*, \**mangV* ‘compact round object’ > *magi*, \**ma-* ‘negative’ > *ma-*, \**niman* ‘louse’ > *iman*, \**ŋaŋa* ‘baby’ > *-ŋaŋ*, \**takVn[V]* ‘moon’ > *takn*

However, \**a* > *e* in the following cases: \**mangat[a]* ‘teeth’ > *meg*, \**sambV* ‘cloud’ > *seb*

- \**i* usually gives *i* (written *y* in some contexts): \**imbi* ‘name’ > *yb*, \**niman* ‘louse’ > *yman*, Proto Madang \**-in* ‘1st singular subject’ > *-in*

- \**u* usually gives *u*: \**kuma-* ‘die’ > *kum-*, \**tumuk/kumut* ‘thunder’ > *tumuk*, \**-un* 1st plural subject’ > *-un*

pTNG \**e* and \**o* are not well attested in Kalam. There is one clear reflex of \**e* and none of \**o*.

- \**[mb, m]elak* ‘light, brightness’ > *melk*

pTNG intervocalic vowels are, under unknown conditions, reflected in Kalam by a short high central vowel [i] which can be analysed synchronically as a non-phonemic transitional vocoid between consonants and so is orthographically zero. Examples: \**[mb, m]elak* ‘light, brightness’ > *melk*, \**kinV-* ‘sleep’ > *kn-*, \**simbi* ‘guts’ > *sb*, \**[si]si* > *ss* ‘urine’.

**Loss of final syllables in pTNG disyllables and trisyllables**

Final syllables (-V, -VC) in pTNG disyllables and trisyllables are sometimes lost in Kalam:

\**imbi* ‘name’ > *yb*, \**simbi* ‘guts’ > *sb*, \**manggat[a]* ‘teeth’ > *meg*, \**apus[i]* ‘grandparent’ > *aps* ‘grandmother’, \**walaka* ‘testicles’ > *walak*

### 3.3 Lexical reconstructions

The following is a list of some 190 lexical reconstructions attributed to pTNG or to later but still early stages. This list comprises most of the reconstructions, given with supporting cognate sets in Pawley (2000a). The items are grouped by fields of meaning. There are about 37 verbs, 10 adjectives, 44 body-part nouns, eight kin terms, six nouns denoting other human statuses, 29 nouns denoting inanimate world things, five terms for artefacts, eight terms for intangible cultural concepts, four terms for insects, seven for birds and bird parts, 11 for plants and plant parts, some 10 forms for independent pronouns, six subject-marking suffixes to verbs, and a few other words.

Of the reconstructions listed, perhaps two-thirds occur in widely distributed subgroups and can be attributed to pTNG with considerable confidence. The rest can be attributed to an early stage of TNG on the grounds that they occur in at least two major subgroups. However, undoubtedly the list includes some items whose distributions have been extended by borrowing.

What makes a set of putative cognates likely to be the outcome of common heritage rather than diffusion? Conformity to regular sound correspondences is one indicator. (However, in many cases the sound correspondences have not been established.) A wide but discontinuous geographic distribution is a second measure. A third indicator is the nature of the meaning(s) represented in the cognate set. Almost all the reconstructions cited here refer to ‘basic’ semantic concepts, denoting body parts (for example head, nose, eye, ear, arm, leg), terms for kinship relations (for example mother, father) and human age-gender status (man, woman, child), salient elements of the inanimate and animate environment (for example rock, water, cloud, moon, ashes, tree, louse, fly, mosquito) and some verbs and adjectives denoting everyday activities (sleep, see/know, hit, low, fly, stand) and important attributes (long, short, cold), and pronouns. Terms for these kinds of concepts are not impervious to borrowing. However, comparative studies around the world have shown that such terms are less likely to be borrowed than terms for culture-specific concepts such as names of domesticated plants, weapons and tools, ornaments, and magico-religious concepts.

#### Some pTNG and early TNG lexical reconstructions organised by semantic fields

C indicates an indeterminate consonant

V indicates an indeterminate vowel

[x] indicates that x may or may not have been present

(x,y) indicates indeterminacy between reconstruction of x and y

|                        |            |             |                   |
|------------------------|------------|-------------|-------------------|
| <b>body parts:</b>     |            | breast      | *amu              |
| arm, forearm           | *mbena     | buttocks    | *simbi + modifier |
| belly, internal organs | *mundun    | cheek       | *mVkem            |
| blood                  | *ke(ñj,s)a | ear         | *ka(nd,t)(i,e)C   |
| bone                   | *kondaC    | excrement 1 | [same as ‘guts’]  |
| brain                  | *muk[V]    | excrement 2 | *ata              |

|                             |                        |
|-----------------------------|------------------------|
| eye 1 (cf. egg 2)           | *(ŋg,k)iti [maŋgV]     |
| eye 2                       | *ŋg(a,u)mu             |
| eye 3                       | *nVpV                  |
| finger nail                 | *(mb,p)utuC            |
| foot, lower leg             | *k(a,o)ond(a,o)C       |
| forehead, head              | *mVtVna                |
| guts, intestines, bowels    | *sim(i,u) or<br>*simbi |
| hair 1                      | *(nd,s)umu(n,t)[V]     |
| hair 2, leaf                | *iti                   |
| hand, claw                  | *sikal or *sakil       |
| head 1                      | *kV(mb,p)utu           |
| head 2                      | *mVtVna                |
| heart 1 (cf. belly, egg 2)  | *mundu-maŋgV           |
| heart 2                     | *simu                  |
| heart 3                     | *kamu                  |
| internal organs (see belly) |                        |
| knee                        | *(ŋg,k)atuk            |
| leg 1                       | *k(a,o)ond(a,o)C       |
| leg 2, calf                 | *kitu                  |
| liver                       | *[ma]pVn               |
| milk, sap                   | *muk                   |
| mouth, teeth                | *maŋgat[a]             |
| navel                       | *simu + modifier       |
| neck 1                      | *k(a,e)(nd,t)ak        |
| neck 2, nape                | *kuma(n,ŋ)[V]          |
| nose                        | *mundu                 |
| penis                       | *mo                    |
| saliva                      | *si(mb,p)at[V]         |
| shoulder                    | *kinV                  |
| skin                        | *(ŋg,k)a(nd,t)apu      |
| teeth (see tooth)           |                        |
| testicles                   | *walaka                |
| tongue 1                    | *mbilaŋ                |
| tongue 2                    | *me(l,n)e              |
| tooth 1 (see mouth)         |                        |
| tooth 2                     | *titi                  |
| urine                       | *[si]si, siti, pisi    |

**kin terms:**

|                         |              |
|-------------------------|--------------|
| father                  | *apa, *mbapa |
| grandparent             | *apus[i]     |
| husband, man            | *ambi        |
| mother, free form       | *am(a,i,u)   |
| mother, bound form      | *na-         |
| sibling, older          | *nan(a,i)    |
| sibling, older same sex | *[mb]amba    |
| sister                  | *aya         |

**age-gender and other social categories:**

|               |                |
|---------------|----------------|
| baby          | *ŋaja          |
| boy           | *nV            |
| man, husband  | *ambi          |
| orphan, widow | *mbenŋa-[masi] |
| woman, female | *panV          |

**birds, bird parts:**

|                    |           |
|--------------------|-----------|
| bird 1             | *n[e]i    |
| bird 2             | *yaka[i]  |
| cassowary          | *ku[y]a   |
| egg 1              | *munaka   |
| egg 2, fruit, seed | *maŋgV    |
| tail               | *a(mb,m)u |
| wing               | *mbutu    |

**insects:**

|           |            |
|-----------|------------|
| butterfly | *apa[pa]ta |
| fly       | *ŋgambu    |
| louse     | *niman     |
| mosquito  | *kasin     |

**plants, plant parts:**

|                         |                |
|-------------------------|----------------|
| bark                    | *ka(nd,t)ap[u] |
| casuarina               | *kal(a,i)pV    |
| fruit, seed (cf. egg 2) | *maŋgV         |
| leaf 1, hair            | *iti           |
| leaf 2                  | *sasak         |
| root                    | *kindil        |
| sap, milk               | *muk           |
| taro                    | *mV            |
| tree, wood              | *inda or *iñja |

**inanimate world:**

|                  |                 |
|------------------|-----------------|
| ashes 1          | *sumbu          |
| ashes 2          | *kambu-sumbu    |
| ashes 3          | *la(ŋg,k)a      |
| cloud 1, sky     | *samb[V]        |
| cloud 2          | *ka(mb,p)utu    |
| fire 1           | *k(a,o)nd(a,u)p |
| fire 2           | *inda           |
| fire 3           | *kambu          |
| flame            | *mbalaŋ         |
| ground 1         | *man[a]         |
| ground 2         | *maka[n]        |
| lightning, light | *(mb,m)elak     |
| moon 1           | *takVn[V]       |
| moon 2           | *kal,(a,i)m     |
| morning          | *k(i,u)tuma + X |

|  |                  |                           |                          |
|--|------------------|---------------------------|--------------------------|
| night  | *k(i,u)tuma      | 1 dual                    | *-uL                     |
| sand   | *sa(ŋg,k)asiŋ    | 2/3 dual                  | *-iL                     |
| sky 1, cloud                                 | *samb[V]         | 1 pl.                     | *-un                     |
| thunder, sky 2                               | *kumut/*tumuk    | 1 sing. different subject | *-pa                     |
| smoke 1                                      | *kambu(s,t)(a,u) |                           |                          |
| smoke 2                                      | *kambu-la(ŋg,k)a | <b>verbs:</b>             |                          |
| stone 1                                      | *kamb(a,u)na     | be (live, stay, sit)      | *mVna-                   |
| stone 2                                      | *[na]muna        | bite                      | *s(i,u)-                 |
| sun 1  | *kamali          | blow                      | *pu + verb               |
| sun 2  | *ketane          | break                     | *pa(ŋg,k)-               |
| water 1                                      | *ok[V]           | burn                      | *nd(a,e,i)-              |
| water 2                                      | *nok             | burn, light a fire        | *ki-                     |
| wind 1                                       | *kumbutu         | carry (on back, shoulder) | *kak(i,u)-               |
| wind 2, breeze                               | *pinVm           | come                      | *me-                     |
|  |                  | cook                      | *adu-                    |
| <b>artefacts:</b>                            |                  | cut, chop                 | *tVk-                    |
| axe  | *tu              | die                       | *kumV-                   |
| fence  | *wati            | do, make                  | *ti-                     |
| netbag 1                                     | *kun             | dream                     | *kina(mb,p)-             |
| netbag 2                                     | *at(i,u)         | eat, drink                | *na-                     |
| string, rope                                 | *asi             | fly, flutter              | *putu(putu) ti-          |
|  |                  | give                      | *mV-                     |
| <b>intangible cultural concepts:</b>         |                  | go 1                      | *pu-                     |
| instructions, language,                      |                  | go 2                      | *yata-                   |
| word, speech                                 | *mana            | hit                       | *[a]tu-                  |
| mind, thought                                | *n(o,u)man       | know, hear, see           | *nVŋg-                   |
| name 1                                       | *imbi            | laugh                     | *ŋgiti (+ verb)          |
| name 2, who                                  | *wani            | lie down, sleep           | *kinV-                   |
| shadow, spirit                               | *k(aw,o)nan      | live, be, sit             | see 'be'                 |
| song, type of                                | *saŋ             | put                       | *(m,p)a(l,t)V-           |
| witchcraft                                   | *koimb, *kum     | say, speak                | *nde-                    |
|  |                  | see, know, perceive       | *nVŋg-                   |
| <b>independent pronouns (for subject,</b>    |                  | shoot                     | *tVmV-                   |
| <b>object, possessor):</b>                   |                  | sleep 1, lie down         | *kinV-                   |
| 1 sing.                                      | *na              | sleep 2                   | *p(e,i)t(e,i)o-          |
| 2 sing.                                      | *ŋga             | speak, talk               | *nde-                    |
| 3 sing.                                      | *ya              | spit                      | *kasipa-                 |
| 3 sing.                                      | *ua              | stand                     | *t(a,e,i)k[V]-           |
| 1 dual                                       | *niLi or *nuLi   | swell                     | *su + verb               |
| 2 dual                                       | *ŋgiLi or *kiLi  | take                      | *(nd,t)a-                |
| 3 dual                                       | *iLi             | tie                       | *ndiŋga-/ndangi-         |
| 1 pl.  | *nu              | turn (oneself)            | *mbuli[ki] + verb        |
| 1 pl.  | *ni              | urinate                   | *X + *si- (urine + verb) |
| 2 pl.  | *ŋgi or *ki      | vomit                     | *mVŋ[g]V ti-             |
| <b>verbal suffixes marking person-number</b> |                  | <b>adjectives:</b>        |                          |
| <b>of subject :</b>                          |                  | blue                      | *muk[V]                  |
| 1 sing.                                      | *-Vn             | cold                      | *kukam(o,u)              |
| 2 sing.                                      | *-an             | dry                       | *ŋgatata                 |

|          |                   |                      |               |
|----------|-------------------|----------------------|---------------|
| full     | *t(o,u)k(i,u) ti- | <b>conjunctions:</b> |               |
| good     | *(nd,t)ebV        | and                  | *ito          |
| heavy    | *kenda            |                      |               |
| long     | *k(o,u)t(u,i)p    | <b>negatives:</b>    |               |
| new      | *kVtak            | not                  | *ma- (+ verb) |
| short    | *tumba            |                      |               |
| straight | *tutu[tu]ku       | <b>numerals:</b>     |               |
|          |                   | two                  | *ta(l,t)(a,e) |

Although about 200 lexical reconstructions are attributable to a very early stage of TNG, no one language today retains more than a small proportion of these. The highest number of reflexes noted so far is in Kalam, a member of the Madang group, with about 40. This relatively high number is probably explained by the fact that there is a good dictionary of Kalam and I happen to be familiar with the language. Even in TNG languages with reasonably good dictionaries sometimes only 20 or so reflexes have been found. In many putative TNG languages for which data are scanty, it is difficult to find more than four or five reflexes. The paucity of cognates is what led Wurm and his associates to conclude that many TNG languages have only a veneer of TNG laid upon an unrelated substrate language.

Now 20 reflexes, or even 50, are not enough to work out in detail the phonological development of a language from pTNG to the present. However all is not lost. This is where bottom-up and middle-level comparisons become essential. Most languages have some quite close relatives, and between them the members of any sizeable subgroup will have many more reflexes than any single language in the group.

### 3.4 Pronouns

The most complete grammatical paradigm so far reconstructed for pTNG is for the independent pronouns. A set of pronoun reconstructions was adumbrated in the 1970s (Wurm 1975a; Voorhoeve 1975), but the pronominal evidence was not systematically tabulated and analysed until Malcolm Ross took up the matter.

This subject is dealt with in some detail in Ross (1995, 2000) and much more briefly in Ross's paper in this volume. Here I will only summarise Ross's reconstructions. Note the pattern whereby the consonant remains constant in the corresponding persons with the singular plural contrast marked by the vowel difference: \*a (singular) vs \*i (plural).

**Table 3:** pTNG free pronouns

|               | 1st person | 2nd person | 3rd person |
|---------------|------------|------------|------------|
| sing.         | na         | ŋga        | [y]a, ua   |
| pl. (i-grade) | ni         | ŋgi        | i          |
| (u-grade)     | nu         |            |            |
| pl.           |            | nja        |            |

We also find evidence for reconstructing dual pronouns. Ross (2000:77, 158–160) refers to a dual suffix \*-li or \*-t and a plural suffix \*-n[V], although he does not say

whether these could be combined with all of the singular pronominal roots or only with some of them. Languages in several widely scattered subgroups have dual pronouns that reflect such a suffix. And among languages that lack a synchronic dual/plural contrast there are some with plural pronouns that appear to reflect old dual forms.

There is support from another quarter for a dual/plural contrast in the free form pronouns. The pTNG free form pronoun roots all have the shape CV or V. All appear ultimately to be cognate with a set of verbal suffixes that mark person-and-number of the subject of the verb and which have the shape VC or V (the metathesis is discussed in section 3.5). These verbal desinences are presented in Tables 4–6 below. The evidence points to pTNG *\*-uL* ‘1st person dual subject’ and *\*-iL* ‘2nd/3rd person dual subject’, where \*L represents a consonant that may have been \*t or \*l or a lateral that was distinct from \*l. There is a formal resemblance to the forms *\*-li* or *\*-t* ‘dual’ that Ross posits as being affixed to free form pronoun roots. Similarly, there is evidence for a verbal suffix *\*-un* ‘1st person plural subject’ that resembles Ross’s plural suffix *\*-n[V]*. The full range of evidence suggests that there were pTNG free form dual pronouns having the approximate forms: *\*niLi* or *\*nuLi* ‘1st plural’, *\*ŋgiLi* or *\*kiLi* ‘2nd dual’ and *\*iLi* ‘3rd dual’.

A number of TNG languages that lack a dual/plural contrast have plural pronouns that appear to reflect such forms.

### 3.5 Verb morphology

#### Subject person-and-number suffixes

It is possible to reconstruct a partial paradigm of suffixes marking subject person-and-number for an early stage of TNG, ancestral to at least several disparate subgroups found in North-East New Guinea, in Madang, Morobe and Eastern Highlands Provinces. The subgroups are (i) Madang, (ii) Finisterre-Huon, and (iii) Goroka-Kainantu. I will refer to the immediate common ancestor of these three groups as Proto North-East New Guinea (pNENG), without implying that such a stage was necessarily distinct from pTNG.

Languages of each of the three subgroups characteristically have several sets of suffixes in independent verbs, each of which (a) mark person-and-number of the subject, (b) generally distinguish singular and plural (and often dual) numbers, and (c) do not distinguish between 2nd and 3rd person suffixes at least in the dual number. Each set is associated with one or more markers of tense, aspect or mood. Some (not necessarily all) tense-aspect-mood (TAM) markers occur as the final or outer suffix on the verb, following the subject suffixes. In some cases the categories of subject and TAM are syncretic, that is are represented by portmanteau suffixes. The phonological interactions between subject and TAM markers, and between suffixes and roots, is one of several factors leading to morphological change and the restructuring of paradigms in TNG languages, complicating the task of rebuilding morphological reconstruction.

Table 4 gives reconstructions of verbal suffixes marking subject in the proto-languages of each of the three groups, based on evidence presented in Pawley (2000b) and Suter (1997).

**Table 4:** Reconstructed verbal suffixes marking subject in Proto Madang, Proto Finisterre-Huon and Proto Kainantu-Goroka Highlands

|              | 1S  | 2S        | 3S        | 1D  | 2D/3D | 1P  | 2P/3P   |
|--------------|-----|-----------|-----------|-----|-------|-----|---------|
| proto Madang | -in | -an<br>-i | -a<br>-an | -uL | -iL   | -un | -ai, -i |
| proto FH     | ?   | -an       | -a, i     | -uL | -iL   | -un | -e, -i  |
| proto KG     | -u  | -an       | -ai, i    | -ur | ?     | -un | -a      |

Most of the proto Madang reconstructions are well supported. There is a pattern in pMadang whereby 1st singular and 1st plural suffixes differ in the vowel, while keeping constant the consonant, *n*. The dual pronouns share a consonant, *t*, while showing a vowel contrast between 1D and 2/3D. Problems lie in the 3S and 2P and 3P forms. *\*-a* is widely reflected as a 3S marker. Only two South Adelbert Range languages show *-i*, but *-i* is common as a 3S marker in certain other TNG groups outside of Madang. Many languages do not distinguish the 2P and 3P suffixes and there is some evidence for reconstructing both *\*-ai* and *\*-i* for 2/3P. However, a fair number distinguish 2P from 3P and there is some evidence for a distinct 2P form, possibly *\*(m)an*.

Our (very tentative) reconstructions for Proto Finisterre-Huon show correspondences with Madang forms in all except 1S. Reconstructed Proto Kainantu-Goroka verbal suffixes show probable cognates with the Madang and Finisterre-Huon paradigms in the 2S, 3S, 1D and 1P forms.

Agreements between the three NENG subgroups cited above allow partial reconstruction of a pNENG paradigm as in Table 5. In the case of the 1S and 2S forms, agreements with TNG languages outside of the three NENG groups, help to resolve indeterminacies.

**Table 5:** pNENG verbal suffixes marking subject

|  | 1S  | 2S  | 3S       | 1D  | 2D/3D | 1P  | 2P        | 3P            |
|--|-----|-----|----------|-----|-------|-----|-----------|---------------|
|  | -Vn | -an | -a<br>-i | -uL | -iL   | -un | -ai<br>-i | -ai<br>-i, -a |

What of TNG groups other than Madang, Finisterre-Huon and Kainantu-Goroka? I have not done a systematic search but for a few of the subject-marking suffixes there are some promising agreements to be found. For example:

|               |                   |  |
|---------------|-------------------|--|
| Angan:        | Baruya            | 1D <i>-olo</i> , 2D <i>-ilo</i> , 1P <i>-ono</i>                                     |
| Chimbu-Wahgi: | Kuman             | 1S <i>-i</i> , 2S <i>-in</i> , 2D <i>-buri</i> , 1P <i>-mun</i> , <i>-umun</i>       |
|               | Salt-Yui          | 1S/1P <i>-i</i> , 2S/2P <i>-n</i> , 1D <i>-bil</i> , 2D <i>-bil</i> , 3D <i>-bil</i> |
|               | Golin             | 1S <i>-bin</i> , 2S <i>-n</i> , 1D/2D/3D <i>-bil</i>                                 |
| Binandere:    | Orokaiva          | 1S <i>-n</i> , 3S <i>-i</i>  |
|               | Korafe            | 1S <i>-n</i>   |
| Suena:        |                   | 1S <i>-n</i> , 3S <i>-i</i>  |
| Dani:         | Grand Valley Dani | 1S <i>-i-</i> , <i>-y</i> , 2S <i>-n</i> , 3S/P <i>-a</i> , 1P <i>-u</i> , <i>-o</i> |

These agreements support the very tentative reconstruction of a partial paradigm for pTNG :

**Table 6:** pTNG verbal suffixes marking subject

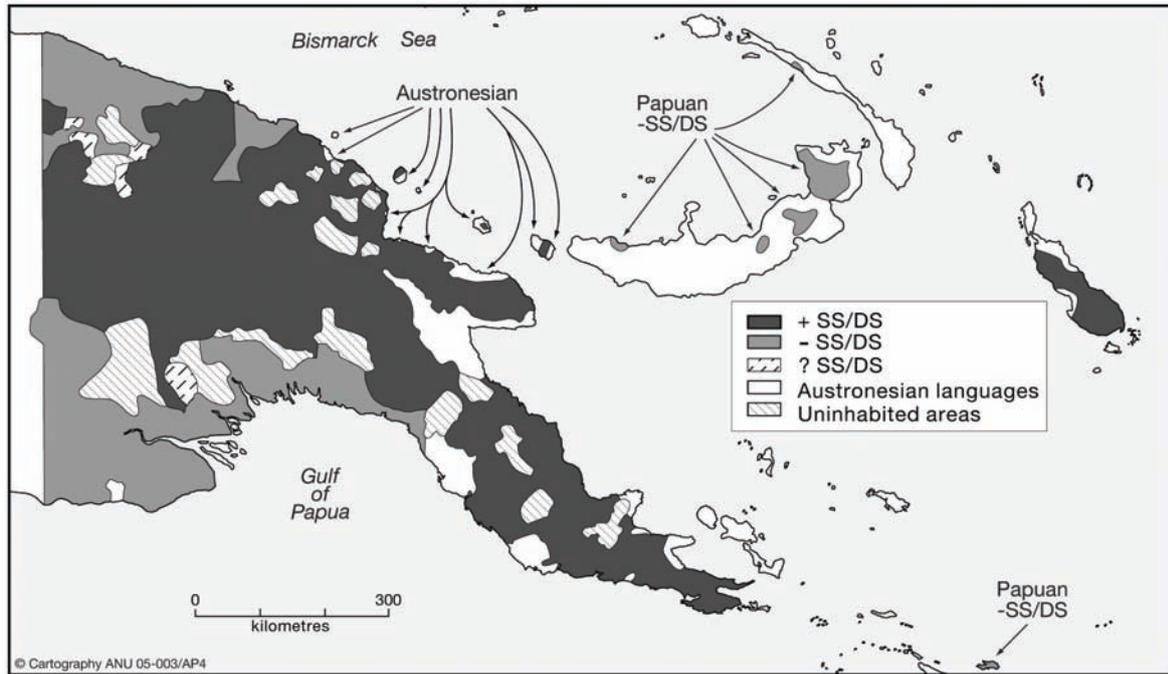
| 1S  | 2S  | 3S     | 1D  | 2/3D | 1P  |
|-----|-----|--------|-----|------|-----|
| -Vn | -Vn | -a, -i | -uL | -iL  | -un |
| -i  |     |        |     |      |     |

Two of the subject-marking suffixes attributed to pMadang show an interesting resemblance to two of the pTNG free form pronouns. pMadang *\*-in* ‘1st singular’ and *\*-un* ‘1st plural’ resemble the pTNG free forms *\*ni* ‘1st plural’ (i-grade) and *\*nu* ‘1st plural’ (u-grade), the formal difference being a metathesis of the consonant and vowel. One might speculate that the pMadang verbal suffixes are ultimately derived from free form pronouns which underwent metathesis after suffixation, with *\*ni* > *\*-in*, and *\*nu* > *\*-un*. In certain contemporary TNG languages we find parallel cases where several of the subject suffixes of one paradigm differ from those of another paradigm in that the order of consonant and vowel are reversed. By extension we may suppose that there were pTNG free form dual pronouns that contained (or consisted of) cognates of the dual subject suffixes but with the order of C and V reversed, that is *\*Lu* ‘1dual’ and *\*Li* ‘2dual’ or ‘2/3 dual’.

**Medial vs final verb morphology**

The most striking cluster of structural features common to most core members of TNG is a morphologically-marked contrast between ‘medial’ and ‘final’ verbs. Final verbs are conventional main verbs, typically carrying suffixes marking absolute tense/aspect/mood and absolute person-number of the subject. Medial verbs are marked for sameness or difference of subject (or in some cases, of pragmatic topic) in relation to the following verb (so-called ‘switch reference’ marking). The medial verb usually carries a suffix marking relative tense; that is, it indicates whether the event denoted by the verb carrying the suffix occurred *before* or *at the same time* as the event denoted by the next verb. (In some languages occurrence after the event denoted by the next verb is a possibility, that is a future or purposive relation). However, TNG languages vary a good deal in the structural details of switch reference marking. Roberts (1988, 1997) has developed a typology of switch reference systems and has carefully charted their distribution within Papua New Guinea (Map 4).

Given their wide distribution across TNG groups, it is very likely that the essential structural distinctions between medial and final verbs outlined above were present in the proto-language. However, this supposition can only be confirmed by discovering sets of cognate suffixes with such functions. Suter (1997) reconstructs *\*-pa* ‘1st person singular different subject’ as a medial verb suffix for an early stage of TNG.



**Map 4:** Distribution of types of switch reference systems in Papua New Guinea (after Roberts 1997)

## 4 On the membership and subgrouping of Trans New Guinea

### 4.1 Membership

In deciding whether a language (or group of languages) belongs to TNG we ask:

- (i) Does it reflect two or more of the pTNG personal pronouns?
- (ii) Does it reflect certain other grammatical paradigms reconstructable for early TNG?
- (iii) Does it continue some of the other basic vocabulary terms reconstructed for pTNG or early TNG?
- (iv) Does it have a sizeable body of cognates exhibiting regular sound correspondences with other TNG languages?

The most comprehensive reassessment of TNG membership is that of Ross (2000), who relies largely on criterion (i), with some attention to the other criteria where evidence permits. It will be recalled that, in TNG III, some 235 languages were assigned to the family as ‘marginal’ members by Wurm et al. (1975). Ross accepts several groups of such languages as members of TNG IV, on the grounds that they meet criterion (i), for example the Papuan languages of Timor, Alor and Pantar, West Bomberai, Gogodola-Suki, Kairi, Marind, and Inanwatan. Ross excludes from TNG a number of languages that were included in TNG III. On the south coast of New Guinea he excludes what he calls the South Central group, comprising Yelmek-Maklew, Morehead-Upper Maro and Pahoturi, and the Eastern Trans Fly group and he reserves judgment on certain languages and groups, including the Eleman and Goilalan groups. Although we provisionally assign the Kiwai group to the TNG phylum on lexical grounds, the evidence at present is slender.

In two areas of mainland New Guinea there is a concentration of small groups and isolates whose TNG status remains unclear. These are (1) the area north of the central highlands from the eastern border of the Geelvink Bay phylum in Irian Jaya to the western part of Sandaun Province in Papua New Guinea (this area contains such problematic groups as Kwerba, Lake Plains-Tor, Border, Nimboran, Sentani and Kaure); (2) the Gulf of Papua area, covering most of the Gulf Province and the adjacent coastal part of the Western Province of PNG. In a recent paper Clouse (1997) has given strong arguments for excluding the Lake Plains languages of Irian Jaya from TNG. Although he suggests, instead, that the Lake Plains languages belong to the Geelvink Bay group, there seems to be no convincing evidence for this alternative (M. Donohue pers. comm.).

## 4.2 Subgrouping

Current understanding of the family tree structure of TNG is uneven. Most of the low-order subgrouping is fairly uncontroversial. Over 40 small groups, typically containing from two to 15 languages, are universally accepted. These small groups are comparable in internal diversity to Germanic or Romance, or to the Polynesian group, and are well-defined, for example Asmat-Kamoro, Awyu-Dumut, Binandere, Chimbu-Wahgi, Dani, Engan, Kalam-Kobon, Koiarian, Mabusu, Mek and Ok. Each group of this order has probably derived from a common proto-language within the last 2000 to 3000 years.

As Foley (1986) predicted, groupings of a high order have proved hard to establish. The following larger subgroups seem reasonably secure. Much of the evidence for these is based on innovations in the personal pronouns, presented in Ross (2000).

### Madang

With about 100 members, Madang is the largest subgroup that can be justified in terms of shared innovations. The most important innovations are the replacement of the pTNG independent pronouns *\*na* 1S, *\*ŋga* 2S and *\*ya*, *\*ua* 3S by Proto Madang *\*ya-*, *\*na-* and *\*nu*, respectively. Other possible innovations defining the group include apparent replacements or irregular changes in several early TNG lexemes: pTNG *\*(n)ok* ‘water’ > Proto Madang *\*yangu*, *\*kiti-mangV* ‘eye’ > *\*ŋgamu*, *\*mbena* ‘arm’ > *\*kambena*, *\*k(a,o)nd(a,o)[C]* ‘leg’ > *\*kani(n)*.

The Madang group corresponds closely to the large lexicostatistically-based group identified by Z’graggen (1975, 1980a-d) as ‘Madang-Adelbert Range’, which he divided into two first order subgroups, each with two main branches. The membership and subgrouping of the Madang group differs from Z’graggen’s account chiefly in the following respects (Ross 2000; Pawley 2005): (1) The Kalam-Kobon group is included in the Madang (it was formerly assigned to Wurm’s East New Guinea Highlands micro-phylum (a group that itself is no longer recognised)). (2) The Madang group is divided into four major branches: (a) South Adelbert Range, (b) Croisilles (roughly corresponding to Z’graggen’s ‘North Adelbert Range’, plus ‘Mabusu’ groups), (c) Rai Coast, and (d) Kalam-Kobon. (3) There is no ‘Brahman’ subgroup: Faita is reassigned to the Josephstaal subgroup of Madang; Biyom and Tauya are reassigned to the Rai Coast subgroup; and Isabi is excluded from Madang and reassigned to the Goroka subgroup.

The extreme structural and lexical diversity found across its major branches suggests that the Madang group probably broke up more than 5000 years ago.

### **Finisterre-Huon**

The Finisterre-Huon group has about 70 member languages. Some share less than 10 per cent of basic vocabulary cognates with each other (McElhanon 1975). Superficial comparison of morphological paradigms suggests that members of the Finisterre-Huon group will on closer study turn out to share several distinctive innovations. Ross (2000) notes a possible innovation in the 1, 2, 3 plural pronouns, all of which end in *-n*, marking plural number.

### **Kainantu-Goroka**

Two well-defined groups are located in the Eastern Highlands, now generally termed the Kainantu and Goroka groups. These have long been tentatively placed together in a subgroup. Ross (2000) reviews previous work and notes three probable innovations that they exhibit in the pronouns: (1) *\*ta[za]* ‘1st pl.’ replaces *\*ni*, *\*nu*, (2) *\*ta-na* ‘2nd pl.’ replacing pTNG *\*ηgi* or *\*ja*, (3) genitive forms ending in *-i*.

### **Chimbu-Wahgi**

This is an uncontroversial group. Ross notes an innovation in the pronoun: the use of *\*im* ‘inclusive marker’, obligatory in 1st inclusive and optional in 2nd and 3rd person plural.

### **South-east Papuan**

Ross finds one piece of evidence for a sizeable South-east Papuan group, consisting of Dagan, Mailuan, Yareba, Manubaran, Kwalean and Koiari but excluding Binandere. This is the replacement of pTNG *\*ηgi* ‘2 PL’ by *\*ya*.

### **Central and South-west New Guinea**

Central and South New Guinea as defined by Voorhoeve (1968) and McElhanon and Voorhoeve (1970) is almost certainly not a valid subgroup. Indeed, Voorhoeve proposed it as a family of (in some cases very remotely) related languages, not as a subgroup. However, we can also be fairly optimistic about validating a sizeable subgroup that includes at least the Ok, Awyu-Dumut and Asmat-Kamoro groups, as suggested by Voorhoeve (1968). I will call this putative group Central and South-west New Guinea.

It should be noted that Ross finds no evidence in the pronouns for such a group. Awyu-Dumut and Asmat do, however, share a rounding of the vowel in the pTNG 1st and 2nd singular pronouns *\*na* and *\*ηga*. It is not clear whether this change is independent or a retention from a common interstage. Ok and Marind both distinguish 3S masculine and feminine: pOk *\*ya*, pMarind *-ye-* ‘3sM’, pOk *\*u-*, pMarind *\*-u-* ‘3sF’. Is this a shared innovation or a retention of an old TNG feature that has been lost in most branches? Ross inclines to the latter view.

### **West Bomberai-Timor**

Ross finds that the West Bomberai and the Timor-Alor-Pantar languages share two probable innovations in pronouns. Both reflect *\*bi* (or *\*ba*) ‘1P’, replacing pTNG *\*ni* or *\*nu*. In West Bomberai the reflexes denote 1st exclusive plural, in contrast to 1st inclusive

plural. In the Timor group the reflexes denote 1st inclusive plural. Both groups also show metathesis of the pTNG independent pronoun *\*na* ‘1s’ to *\*an*, although Mark Donohue (pers. comm.) points out the possibility that *\*an* was the earlier form of the pronoun, preserved in the subject suffixes discussed in section 3.5.

## **5 Culture historical notes**

### **5.1 Introduction**

There remains the question: What culture-historical implications can be drawn from the linguistic evidence about the history of TNG languages, or from a conjunction of linguistic and other lines of evidence? In particular, where and when was pTNG spoken? And what circumstances—technological, demographic, geographic—led to the spread of TNG languages?

### **5.2 On the location of proto TNG and directions of dispersal**

The linguistic diversity of the New Guinea area (New Guinea and nearby smaller islands) is extraordinary—unequaled in the world by any other region of comparable size. Not only does the New Guinea area contain over 1000 languages, but these fall into more than 20 families not known to be related to each other. This order of diversity is consistent with archaeological evidence showing that human settlement of New Guinea, New Britain and New Ireland goes back at least 40,000 years and that Bougainville was settled at least 28,000 years ago (Allen & Gosden 1996; Pavlides & Gosden 1994; Specht this volume; Spriggs 1997). It appears that only one of the language families is a recent arrival: Austronesian. Archaeological dates for the first appearance of Austronesian-associated assemblages in north-west Melanesia are around 1500–1300 BC (Bellwood 1997; Spriggs 1997; Kirch 1997).

The TNG family is exceptional among the families of Papuan languages in its large membership and wide geographic spread. TNG is predominantly a family of the mountain ranges that run along the centre of New Guinea. It has branches in certain parts of the lowlands of New Guinea and in the Timor area but it is likely that all of these branches, ultimately, represent expansions from the highlands.

It is likely that the pockets of Papuan languages in New Guinea that do not belong to the TNG family are linguistic residual areas. That is to say, these languages belong to ‘old’ families which have managed to resist the TNG expansion. As Map 4 shows, the non-TNG families and isolates are all quite small and localised. In Papua New Guinea most of them are located to the north of the central cordillera, especially in three provinces: Sandaun, Sepik and Madang. In Irian Jaya most of them are located on the Bird’s Head and, east of the Bird’s Head, to the north of the main highlands range, from Cenderawasih Bay to the Mamberamo River.

In Island Melanesia it is clear that the pockets of Papuan languages that survive—chiefly in New Britain, Bougainville and the central Solomons—are the residue of a once much more extensive body of languages that were spoken in this region before the arrival and expansion of Austronesian languages over the last three millennia or so (Ross 2001; Dunn et al. 2002).

Can we say any more than that pTNG was probably spoken somewhere in the chain of mountains that runs down the centre of New Guinea? Measured in terms of the density of currently established high-order subgroups of TNG, the region of greatest diversity is that area of Papua New Guinea between the Strickland River in the west and the Eastern Highlands Province in the east, together with Madang Province and the Finisterre Ranges and Huon Peninsula Province to the north. It is safe to say that this was a very early area of TNG expansion. Whether it was the *original* dispersal centre is another matter. It is true that the highlands in Irian Jaya contains fewer high-order subgroups. That is not to say, however, that TNG languages have not been present in this region for as long, or almost as long as they have been in the eastern highlands. Without a better understanding of the high-order subgrouping than exists for the family at present we cannot on linguistic grounds alone confidently identify its primary dispersal centre. However, there are other relevant lines of evidence.

### 5.3 On dates of dispersal

When did Proto TNG break up? The main kinds of evidence that bear on this include:

- (1) degree of diversity in basic vocabulary and morphology within and between the subgroups of TNG languages;
- (2) dating of certain archaeological or geomorphical events that can be correlated with linguistic events; and
- (3) relative chronologies indicated by borrowing in relation to subgroups.

#### Lexical diversity

Degree of diversity in basic vocabulary is a very crude indicator of the age of a language family, that is of the date when its immediate common ancestor split up. It is instructive to compare the diversity within TNG to that of Indo-European and Austronesian, two families whose chronologies are fairly well established. Lexical diversity within TNG is far greater than within either Indo-European or Austronesian. Indeed the largest subgroup of TNG so far identified, Madang, probably exhibits greater lexical diversity than either of those families. As was noted earlier, geographically widely separated subgroups of TNG generally show less than 7 per cent of cognates in basic vocabulary and in some cases only 2–3 per cent.

Austronesian is generally thought to have broken up by at least the late 3rd millennium BC, by which time the neolithic culture associated with Austronesian languages was widely dispersed around Taiwan and had been carried into the Philippines (Bellwood 1997, 2000; Pawley 2002). Blust (1999) argues that Austronesian consists of some 10 primary branches of which nine are spoken on Taiwan, with all members of the family spoken outside Taiwan forming a single subgroup, Malayo-Polynesian. The Taiwan languages generally share between 15 and 25 per cent of basic vocabulary with the Malayo-Polynesian languages of the Philippines and Indonesia.

There are two schools of thought regarding the dates at which Proto Indo-European broke up. The dominant view at least until recently, has been that Indo-European probably broke up not later than the 4th millennium BC. The major subgroups of Indo-European (IE)—Celtic, Romance, Germanic, Slavic, Indo-Iranian, and so on—generally converge at

around 20 per cent of cognates in the 200-item list. Recently there has been some support for a view that while a date of around 4000 BC may suffice for the major North European expansion, the initial spread of Indo-European began as early as about 7000–8000 BC, from Anatolia (Renfrew 1987; Gray & Atkinson 2003). This view remains a minority one among Indo-Europeanists.

Attempts to develop an absolute-dating method, glottochronology, from lexicostatistical studies have been fraught with problems (Renfrew et al. eds 2000). If we were to trust glottochronology we would conclude that pTNG broke up several millennia earlier than either proto Indo-European or proto Austronesian, perhaps as early as 10 millennia ago.

Substantial variations in the lexical retention rates of European languages have been documented. Blust (2000) finds that this is also the case for Austronesian languages. However, by using generally accepted subgroupings and lexical reconstructions as a baseline, it is possible to identify those Indo-European and Austronesian languages that have unusually high retention rates (lexically conservative languages) or unusually low retention rates (innovative languages) and those that cluster around a median rate. Glottochronological dates for splits within this middle group are likely to be more reliable than those for the comparisons involving high and low retention rate languages.

The question arises whether all members of the TNG phylum have a low retention rate in basic vocabulary or whether TNG shows a range of variation similar to Austronesian or Indo-European. It seems unlikely that *all* TNG languages would have replaced basic vocabulary faster than *all* Austronesian or Indo-European languages. I conclude that TNG is probably an older family than either Austronesian or Indo-European.

#### **5.4 Was the TNG expansion powered by agriculture?**

Was the initial spread of TNG languages powered by new technology and associated changes in settlement patterns and demography? It is a reasonable supposition to link early highlands agricultural systems with TNG speakers.<sup>6</sup> It seems unlikely that the TNG family would have achieved its present remarkable distribution unless its speakers held some cultural advantages that enabled them to build up populations that could (a) expand fairly rapidly along the central cordillera of New Guinea, and (b) maintain continuous habitation of the major highlands valleys, through periods of change in climate, vegetation and fauna.

At present, however, the connection between agriculture and the TNG expansion is no more than circumstantial. We lack linguistic evidence that directly points to knowledge of farming by speakers of pTNG or early stages of TNG. At this stage a term for ‘taro’ (something like *\*ma*) is about the only relevant lexical reconstruction that can be tentatively attributed to early TNG because of its wide distribution. However, because the term *\*ma* ‘taro’ stands alone, instead of being embedded in a full terminology for parts of the plant and practices associated with its cultivation, diffusion cannot be ruled out. I know of no other widely distributed cognate sets for names of plants and their parts and for implements and processes associated with their cultivation. There are some names for plants and other entities that have spread recently but these do not count in the search for early TNG words.

---

<sup>6</sup> At the *Perspectives on the Bird's Head* conference a few years ago I suggested such a link. (Pawley 1998:684).

Most Pacific prehistorians favour a gradual model of plant domestication in Near Oceania (north-west Melanesia) and Indonesia, beginning with the semi-domestication of wild species in the late Pleistocene. It is now widely held that arboriculture, as opposed to agriculture, began in this region more than 10 millennia ago, with the planting and transportation of certain trees that bear edible fruit, especially various species of the genus *Canarium* but possibly also *Artocarpus*, *Pandanus brosimus*, *P. julianettii* and *P. conoidea*, *Cordyline fruticosa*, *Dracanotomelon*, *Inocarpus*, *Spondias*, *Burckella* and others (Kennedy & Clarke 2004; Yen 1991; Spriggs 1997).

The central highland valleys of New Guinea were visited by humans, although probably not occupied continuously, during the late Pleistocene. The climate was much colder during the late Pleistocene than now and the treeline was much lower but warmed up significantly after 14ka BP. Hunter-gatherers evidently exploited pandanus nuts and hunted megafauna of the grasslands and perhaps of the mountain forest, probably on a seasonal basis. In the central highlands, landscapes begin to be modified by humans at a number of sites after 10,000 with a marked increase from about 5000 years ago (Hope & Haberle this volume). There is increasingly strong evidence for some form of agriculture as early as 10,000 BP in the Upper Wahgi Valley (Golson 1977; Golson et al. in prep.; Denham this volume; Denham et al. 2003). The earliest phase of the Kuk swamp sequence, dated to around 10,000 BP, indicates shifting cultivation on the wetland edge, with pits, stakeholes, postholes, runnels, consistent with planting and tethering. At that time the Kuk site seems to have presented a more favourable environment for habitation than other Highland valleys. Denham et al. (2003:190) write that

Unlike other valleys in the uplands, the grasslands within the Kuk swamp catchment did not succumb to the forest advance at the onset of the Holocene. Instead, the grasslands and fern flora increased at the expense of forest between 10,200 and 7400 cal yr under the influence of periodic fire episodes and probably anthropogenic clearance.

The main cultivated plants are thought to have been *Colocasia* taro and bananas. Starch grains of *Colocasia esculenta* and *Musa* bananas are present as phaeoliths. Taro is a lowlands plant but Denham (2002) argues that it had spread naturally into the Highlands by 10,000 BP.

In phase 2, dated to 6900–6400, there was mounding and draining of wetland soils, consistent with intensive cultivation, implying a high dependence on food production relative to foraging. A higher incidence of taro and bananas remains are present, and in a grassland environment it is unlikely that bananas would have grown wild in such frequency. In phase 3, dated to 4350–3980 BP, there are sequential ditch networks linking major drainage channels.

The shift from a foraging to a primarily agricultural economy may have taken place over many millennia at Kuk. As to how fast and far agriculture spread in New Guinea in the period 10,000 to 3000 BP, the archaeological evidence at present says little. There are several sites in the Upper Wahgi Valley with well-dated drainage systems older than 3000 BP (Denham 2002). These remain the only New Guinea sites of this kind with secure dates although there is another early site at Yeni swamp in the Lower Jimi Valley with signs of drainage structures at 5000 BP (Gillieson, Gorecki & Hope 1985). Pollen analysis shows that reduction in forests due to burning had also taken place in the Kelala swamp in the Baliem Valley by 7800 BP, although in the Tari Basin in the Southern Highlands of Papua New Guinea it is first evident only at 1700 BP (Hope & Golson 1995; Hope & Haberle

this volume). There is as yet no direct evidence that such burning was associated with agriculture. However, in the case of the Baliem Valley, Golson observes that pollen cores record an almost

continuous vegetation history from beyond 7000 bp to the present, reflecting progressive human impact by way of agriculture through the increasing representation of secondary forest taxa and associated changes ... This new evidence from the Baliem is the strongest independent support for the claims of 9000 year old agriculture based on Kuk. (Golson 1991:487)

There is as yet no archaeological record of agriculture in the New Guinea lowlands before 3500 BP, where few early and mid-Holocene sites have been found, or in Island Melanesia.

Not all New Guinea societies were farmers even in historic times. Roscoe (this volume) argues that more than 30 lowland societies at first contact were primarily foraging, relying heavily on wild stands of sago as their main source of carbohydrates. And in some New Guinea agricultural societies hunting and gathering continued to provide very significant supplementary food sources in recent times.

None of the domestic animals that were important in Near Oceania at first contact—pigs, dogs and chickens—were native to the region. They were however all part of the Austronesian cultural package in Island Southeast Asia. There has been vigorous debate over the antiquity of the pig in New Guinea. Did it predate the arrival of AN speakers? Bulmer (1975, 1982, 1998) and Allen (1993) report evidence of pig teeth from several sites associated with pre-Lapita dates. Others, such as Spriggs (1997) and Bellwood (1997), think the evidence is unconvincing. A parallel debate has taken place over the antiquity of pottery in New Guinea. Pottery is made among many Papuan-speaking communities in the Sepik–Ramu region as well as in Oro Province in south-east New Guinea. Pottery sherds in archaeological sites in the Sepik–Ramu and Simbai areas that may predate the Lapita horizon are reported by Bulmer (1982), Gorecki (1991), and Swadling et al. (1989). The dating of these materials remains controversial (Spriggs 1997).

Where the shift to intensive agriculture did occur it must have brought radical changes in patterns of social organisation and material culture. Agriculturalists are sedentary, tied to the land they have cleared, tilled, planted and fallowed. There is potential for faster population growth, larger social units and social hierarchy and for the making of ‘heavy’ artefacts, such as substantial houses, elaborate carvings and large containers. Language populations tend to become larger and this in turn must have allowed more marriage within the language community. Ethnographic evidence suggests that the shift to intensive agriculture occurred faster in certain regions than others—the broad, fertile highland valley floors being among the first.

At this stage it is not clear how far comparative work will enable lexical reconstructions to extend into the domain of ‘material culture’ for early stages of TNG, including the cultivation of plants. As yet no historical linguist has undertaken a thorough, New Guinea-wide search for cognates in cultural domains (however, see the work of an anthropologist, Hays, this volume). The job is made harder by major gaps in the descriptive sources.<sup>7</sup> It is also slowed by the lack of manpower. There is not a single linguist whose primary research field is Papuan historical linguistics. Only a handful of linguists are active in

---

<sup>7</sup> Apart from manpower, the lack of good dictionaries remains the central problem in Papuan historical linguistics.

TNG historical studies. It might be said that studies of the Trans New Guinea family are about where Indo-European studies were in the 1820s, in the days of Rask and Grimm, but with the prospect of having only a tiny fraction of the manpower that was available for the study of Indo-European.

### **Geographic constraints**

The possible directions of spread of TNG languages have been constrained by a number of geographic factors, including sea level shifts. Since the height of the last Ice Age, around 21,000 years ago, the coastlines of New Guinea, particularly on the southern side, have changed a good deal. Chappell (this volume) gives an up-to-date assessment of the major changes. After 21 ka the Sahul-Arafura shelf which linked Australia and New Guinea was gradually flooded by rising seas, with the last land connections (through Torres Straits) inundated shortly before 8 ka. By 6 ka rapid changes to the New Guinea coastline ceased, with changes since then largely confined to the progradation of coastal plains and deltas.

However, Chappell refers to significant gradual changes in two regions over the last six millennia. During the mid-Holocene much of what is now the Sepik–Ramu coastal plain was below sea level. A shallow sea extended inland, at its maximum possibly as far as Ambunti (Swadling 1997) and almost certainly west of the confluence of the Yuat and Sepik. According to Chappell the inland sea reached its maximum extent about 5500–6500 BP, then contracted gradually under deposits of alluvial soils from the two river systems. This inland sea would have separated the central highlands from higher-lying areas of north-east New Guinea in what are now the Sepik and Sandaun provinces. It is noteworthy that TNG languages are largely absent from these areas.

The other region showing significant coastline changes during the last few millennia is part of the extensive Fly-Digul Platform in southern New Guinea (Chappell this volume). The low-lying Digul River region, which forms the western half of this platform, was invaded by the sea and inundated about 6000 years ago, as was the delta and narrow floodplain of the Fly River. It appears that most of the swampy Digul lowland has been gradually established over the past six millennia as a result of sedimentary deposits from the southwards-flowing rivers. The western half of the Fly-Digul Platform roughly corresponds to the area now inhabited by the Asmat and Kamoro peoples, who speak seven closely related languages (Voorhoeve 2001, this volume). From this conjunction of facts—the relative homogeneity of Asmat Kamoro and the changes in the geography of the coastline of south-west New Guinea—we can infer that the expansion of Asmat-Kamoro languages across their present territory is a fairly recent event, probably occurring within the past two or three millennia. Voorhoeve (this volume) suggests that before they moved down to the southern plains, speakers of the ancestral Asmat-Kamoro language once lived in the highlands of Irian Jaya, possibly around the headwaters of the Sirac River. He concluded that some of the lexical resemblances between Asmat-Kamoro and the Ok languages are due to ancient contact, contact that can only have occurred in the mountains.

It is uncertain whether the Timor-Alor-Pantar branch of TNG was well established in the Timor region before Austronesian speakers arrived there at least 3000 years ago. My impression is that the Timor-Alor-Pantar languages are lexicostatically more diverse than the Austronesian languages of the same region but systematic comparisons have yet to be carried out.

## **Borrowing events**

Austronesian loanwords in TNG languages are a possible indicator of the age of some TNG subgroups. Austronesian speakers first moved into north-west Melanesia about 3500–3300 BP. It follows that if Austronesian loanwords were already present in a certain TNG interstage (an intermediate proto-language), that interstage must postdate contact with Austronesian speakers. It is clear that many of the Austronesian languages now present on the New Guinea mainland arrived only within the past 2000 years. However, there is some evidence (Ross 1988:21) of lexical borrowing from Austronesian sources by TNG languages of the north coast of New Guinea that predates these more recent arrivals. What dates the loans as early is that they retain original Oceanic root final consonants, whereas these consonants have been lost in all contemporary Oceanic languages of the north coast of New Guinea. Loans showing such final consonant retentions appear to be restricted to certain branches of the Madang subgroup of TNG.

McElhanon and Voorhoeve (1970) and Lynch (1981) pointed to instances of possible borrowings from Austronesian languages, including items of basic vocabulary, that are more or less widespread among TNG languages. However, on closer study the case for Austronesian borrowings in TNG basic vocabulary appears to be weak, as Chowning (1987) has argued with some force. There are however a few clear cases of borrowing in cultural vocabulary, for example reflexes of POc *\*boRok* ‘pig’ are found widely in Papuan languages, while reflexes of pTNG *\*mV* ‘taro’ are found in a number of Austronesian languages.

## **6 Conclusions**

The directions and chronology of the TNG dispersal, and its technological and demographic concomitants, remain somewhat shadowy. However, it is at least encouraging that there is some measure of agreement between the testimonies of historical linguistics, archaeology, palaeobotany and palaeogeography on some of these matters.

The following observations are put forward.

- (1) The wide geographic spread of TNG languages, from one end of New Guinea to the other (and into the Timor area) stands in sharp contrast to all other ‘Papuan’ language families. The spread of TNG languages was recent enough for their common origins to be still detectable, yet early enough for the language family to be far more diverse than either Indo-European or Austronesian. These two facts indicate that pTNG broke up no later than about 6000 BP and possibly as early as 10,000 BP.
- (2) TNG is predominantly a family of the central mountain ranges of New Guinea. In particular most of its high-order subgroups of TNG are concentrated in the central highlands. Such a distribution indicates that speakers of TNG languages have occupied the large valleys of the central highlands continuously for several millennia. This continuity suggests that, at least in these large valleys, there were quite sizeable stable populations.
- (3) There is archaeological and palynological evidence of agriculture in the central highlands as early as 10,000 BP and evidence of intensification of agriculture in the Wahgi Valley, and probably in the Baliem Valley by about 6000 BP. It is reasonable to assume that an early branch of TNG speakers was established in the Wahgi Valley

by at least 6000 BP, and (less certainly) that another was established in the Baliem Valley by a similar date. If both were true it would imply that some of the intervening valleys were also occupied by TNG speakers by that time.

- (4) Six thousand years ago a shallow inland sea occupied much of what is now the flood plain of the middle and lower Sepik and Ramu rivers. To the north of this shallow sea lay a broad band a coastal ranges and coastal plains, separated by it from the central highlands. It is noteworthy that several non-TNG language families are today found in that region and in the new low-lying land gradually created from the Sepik and Ramu sediments. The inland sea may for a time have been a barrier limiting the advance of TNG speakers into the Sepik and Ramu basins. Malaria may also have been a major demographic constraint.
- (5) However, TNG speakers did successfully establish themselves in the swampy lowlands of the Digul River basin and nearby coastal plains of SW New Guinea, probably within the last 3000 years or so.
- (6) TNG languages are well represented in the Huon Peninsula and certain other parts of the north-east coast of New Guinea. Although the Vitiaz Strait, separating New Britain from the Huon Peninsula, is not a very formidable ocean crossing, no TNG languages are found in the Bismarck Archipelago.<sup>8</sup>
- (7) It appears that TNG speakers reached the Timor-Alor-Pantar area several millennia ago. The question arises, did they get there before the arrival of Austronesian speakers some 3500 years ago? Careful study of the internal diversity of the TNG language in the Timor region may help to answer this question.

## References

- Allen, Jim, 1993, Notions of the Pleistocene in Greater Australia. In M.A. Smith, M. Spriggs and B. Fankhauser, eds *Sahul in review: Pleistocene archaeology in Australia, New Guinea and Island Melanesia*, 139–151. Canberra: Department of Prehistory, Australian National University.
- Allen, Jim and Chris Gosden, 1996, Spheres of interaction and integration: modelling the culture history of the Bismarck Archipelago. In Davidson et al., 1996:183–197.
- Baak, Connie, Mary Bakker and Dick van der Meij, eds, 1995, *Tales from a concave world. Liber Amoricum Bert Vooerhoeve*. Leiden University, Projects Division, Department of Languages and Cultures of South-East Asia and Oceania.
- Bellwood, Peter, 1997, *Prehistory of the Indo-Malaysian Archipelago*. Honolulu: University of Hawaii Press.
- 2000, The time depth of major language families: an archaeologist's perspective. In C. Renfrew, A. McMahon and L. Trask, eds *Time depth in historical linguistics*, vol. 1, 109–140. Cambridge: McDonald Institute for Archaeological Research.

---

<sup>8</sup> One TNG language, Kovai, is spoken on Umboi. It is not clear how or when (the ancestor of) this language reached Umboi. Kovai is a member of the Huon branch of the Finisterre-Huon subgroup. The lexical distance of Kovai from other Huon languages suggests that it has been separated from them for at least two or three millennia.

- Bellwood, Peter and Colin Renfrew, eds, 2002, *Examining the farming/language dispersal hypothesis*. Cambridge: McDonald Institute for Archaeological Research.
- Blust, Robert A., 1999, Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. In E. Zeitoun and P.J-K Li, eds *Selected papers from the Eighth International Conference on Austronesian Linguistics*, 31–94. Taipei: Academia Sinica.
- 2000, Why lexicostatistics doesn't work: the 'universal constant' hypothesis and the Austronesian languages. In Renfrew et al., eds 2000:311–332.
- Bowern, Claire and Harold Koch, eds, 2003, *Australian languages: classification and the comparative method*. Amsterdam/Philadelphia: John Benjamins.
- Bulmer, Susan, 1975, Settlement and economy in prehistoric Papua New Guinea; a review of the archaeological evidence. *J. Société des Oceanistes* 31:7–75.
- 1982, Human ecology and cultural variation in prehistoric New Guinea. *Monographiae Biologicae* 42:169–206. The Hague: Dr W. Junk.
- Capell, Arthur C., 1948–1949, Distribution of the languages in the Central Highlands, New Guinea. *Oceania* 19(2–4):104–129, 234–253, 349–377.
- Chowning, Ann, 1987, The supposed Austronesian content of the Trans-New Guinea Phylum. In D.C. Laycock and W. Winter, eds *A world of language: papers presented to Professor S.A. Wurm on his 65th birthday*, 119–125. Canberra: Pacific Linguistics.
- Clouse, Duane A. 1997, Towards a reconstruction and reclassification of the Lakes Plain languages. In K. Franklin, ed. *Papers in Papuan linguistics* No. 2. 133–236. Canberra: Pacific Linguistics.
- Comrie, Bernard A., 1986, Haruai verb structure and language classification in the Upper Yuat. *Language and Linguistics in Melanesia* 17:140–160.
- 1989, Genetic classification, contact, and variation. In T.J. Walsh, ed. *Synchronic and diachronic approaches to linguistic variation and language change*. Georgetown University Round Table on Languages and Linguistics 1988, 81–93. Washington, D.C.: Georgetown University Press.
- Denham, Tim, 2002, Archaeological evidence for mid-Holocene agriculture in interior New Guinea: a critical review. *Archaeology in Oceania* 38(3):159–176.
- Denham, T.P., S.G. Haberle, C. Lentfer, T. Fullagar, J. Field, M. Therin, N. Porch and B. Winsborough, 2003, Origins of agriculture at Kuk Swamp in the Highlands of New Guinea. *Science* 201:189–193.
- Dixon, R.M.W., 2002, *The Australian languages*. Cambridge: Cambridge University Press.
- Dunn, Michael, Ger Reesink and Angela Terrill, 2002, The East Papuan languages: a preliminary typological appraisal. *Oceanic Linguistics* 41:28–62.
- Evans, Nicholas, ed., 2003, *The non-Pama-Nyungan languages of Australia: comparative studies of the continent's most historically complex region*. Canberra: Pacific Linguistics.
- Foley, William A., 1986, *The Papuan languages of New Guinea*. Cambridge University Press.
- 2000, The languages of New Guinea. *Annual Review of Anthropology* 357–404.
- Franklin, Karl J., 1975, Comments on Proto-Engan. In S.A. Wurm, ed. 1975:263–275.

- Gillieson, D., P. Gorecki and G. Hope, 1985, Prehistoric agricultural systems in a lowland swamp, Papua New Guinea. *Archaeology in Oceania* 20:32–37.
- Golson, Jack, 1977, No room at the top: agricultural intensification in the New Guinea Highlands. In J. Allen, J. Golson, R. Jones, eds *Sunda and Sahul: prehistoric studies in Southeast Asia, Melanesia and Australia*, 601–638. London & New York: Academic Press.
- 1991, Bulmer phase II: Early agriculture in the New Guinea Highlands. In A. Pawley, ed. 1991:484–491.
- Golson, J., T. Denham, P. Swadling and J. Muke, in prep., *Ten thousand years of gardening: Kuk and the archaeology of agriculture in Papua New Guinea*.
- Greenberg, Joseph H., 1971, The Indo-Pacific hypothesis. In Thomas A. Sebeok, ed. *Current trends in linguistics 8: linguistics in Oceania*, 807–871. The Hague: Mouton.
- Gray, Russell and Quentin Atkinson, 2003, Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426:435–439.
- Haiman, John, 1979, Review of Wurm (ed.) 1975. *Language* 55(4):894–903.
- Healey, Alan, 1964, The Ok language family in New Guinea. PhD thesis, Australian National University.
- Heeschen, Volker, 1978, Review of Wurm (ed.) 1975. *Anthropos* 72:978–980.
- Kennedy, Jean and William Clarke, 2004, *Cultivated landscapes of the southwest Pacific*. RMAP Working Paper No. 50. Research School of Pacific and Asian Studies, The Australian National University.
- Kirch, Patrick V., 1997, *The Lapita peoples. Ancestors of the Oceanic world*. Oxford: Blackwell.
- Lang, Ranier, 1976, Review of Wurm (ed.) 1975, *Kivung*, *Journal of the Linguistic Society of Papua New Guinea* 9(1):72–80.
- Lynch, John, 1981, Austronesian loanwords (?) in Trans-New Guinea Phylum vocabulary. *Papers in New Guinea linguistics*, No. 21, 165–180. Series A-61. Canberra: Pacific Linguistics.
- McElhanon, Kenneth A., 1967, Preliminary observations on Huon Pensinsula languages. *Oceanic Linguistics* 6:1–45.
- 1970, Lexicostatistics and the classification of the Huon Peninsula languages. *Oceania* 40(3):214–231.
- 1975, North-Eastern Trans New Guinea Phylum languages. In Wurm, ed. 1975:527–567.
- McElhanon, K. and C.L. Voorhoeve, 1970, *The Trans-New Guinea Phylum: explorations in deep-level genetic relationships*. Canberra: Pacific Linguistics.
- Pavrides, Christina and Chris Gosden, 1994, 35000 year-old sites in the rainforests of West New Britain, Papua New Guinea. *Antiquity* 68:604–610.
- Pawley, Andrew, 1995, C.L. Voorhoeve and the Trans New Guinea Phylum hypothesis. In Baak et al. 1995:83–122.
- 1998, The Trans New Guinea Phylum hypothesis: a reassessment. In Jelle Miedema, Cecilia Ode, Rien A.C. Dam, eds *Perspectives on the Bird's Head of Irian Jaya, Indonesia*, 655–689. Amsterdam: Editions Rodopi.

- 2000a, Some Trans New Guinea phylum cognate sets. Printout. Department of Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- 2000b, Notes on early Trans New Guinea verb morphology: reconstructing subject suffixes. Printout. Department of Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- 2001, The Proto Trans New Guinea obstruents: arguments from top-down reconstruction. In A. Pawley, M. Ross and D. Tryon, eds 2001:261–300.
- 2002, The Austronesian dispersal: people, language, technology. In Bellwood and Renfrew, eds *Examining the language/farming dispersal hypothesis*, 251–273. Cambridge: McDonald Institute for Archaeological Research.
- 2005, Madang languages. In Keith Brown, ed. *The encyclopaedia of language and linguistics* (2nd edition) vol. 7, pp. 429–432. Oxford: Elsevier.
- in prep, The Proto Trans New Guinea resonants. Department of Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- Pawley, Andrew, ed., 1991, *Man and a half: essays in Pacific anthropology and ethnobiology in honour of Ralph Bulmer*. Auckland: Polynesian Society.
- Pawley, Andrew, M. Ross and D. Tryon, eds, 2001, *The boy from Bundaberg: studies in Melanesian linguistics in honour of Tom Dutton*. Canberra: Pacific Linguistics.
- Renfrew, Colin, 1987, *Archaeology and language*. London: Jonathan Cape.
- Renfrew, Colin, April McMahon and Larry Trask, eds, 2000, *Time depth in historical linguistics*. Cambridge: MacDonal Institute for Archaeological Research.
- Roberts, John R., 1988, Switch-reference in Papua New Guinea languages: a syntactic or extra-syntactic device? *Australian Journal of Linguistics* 8:75–118.
- 1997, Switch-reference in Papua New Guinea: a preliminary survey. In A. Pawley, ed. *Papers in Papuan linguistics*, 101–241. Canberra: Pacific Linguistics.
- Ross, Malcolm D., 1988, *Proto Oceanic and the Austronesian languages of Western Melanesia*. Canberra: Pacific Linguistics.
- 1995, The great Papuan pronoun hunt: recalibrating our sights. In Baak et al. 1995:139–168.
- 2000, Pronouns as preliminary evidence for grouping Papuan languages. Printout, Department of Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- 2001, Is there an East Papuan phylum? Evidence from pronouns. In A. Pawley, M. Ross and D. Tryon, eds 2001:301–321.
- Shaw, Daniel, 1986, The Bosavi language family. Series A-70, 45–76. Canberra: Pacific Linguistics.
- Spriggs, Matthew, 1997, *The Island Melanesians*. Oxford: Blackwell.
- Suter, Edgar, 1997, A comparative look at the dual and plural forms of inflections and pronouns in Northeast New Guinea Papuan languages. *Language and Linguistics in Melanesia* 28:17–68.

- Swadling, Pamela, 1997, Changing shorelines and cultural orientations in the Sepik–Ramu, Papua New Guinea: implications for Pacific prehistory. *World Archaeology* 29:1–14.
- Swadling, Pamela, J. Chappell, G. Francis, N. Araho and B. Ivuyo, 1989, A late quaternary inland sea and early pottery in Papua New Guinea. *Archaeology in Oceania* 24:106–109.
- Voegelin, Carl F. and Florence M. Voegelin, 1977, *Classification and index of the world's languages*. New York/Amsterdam: Elsevier.
- Voorhoeve, C.L., 1968, The Central and South New Guinea Phylum: a report on the language situation in south New Guinea. In *Papers in New Guinea linguistics* No. 8, 1–17. Canberra: Pacific Linguistics.
- 1975, Central and Western Trans-New Guinea Phylum languages. In Wurm, ed. 1975:344–459.
- 1987, Worming one's way through New Guinea: the chase of the peripatetic pronouns. In Donald C. Laycock and Werner Winter, eds *A world of language: papers presented to Professor S.A. Wurm on his 65th birthday*, 709–727. Canberra: Pacific Linguistics.
- 2001, Proto Awyu-Dumut phonology II. In Pawley, Ross and Tryon, eds, 2001:345–360.
- Wurm, Stephen A., 1960, The changing linguistic picture of New Guinea. *Oceania* 31(2):121–136.
- 1964, Australian New Guinea Highland languages and the distribution of their typological features. *American Anthropologist* 66(4/2):79–97.
- 1965, Recent comparative and typological studies in Papuan languages of Australian New Guinea. *Lingua* 15:373–399.
- 1971, The Papuan linguistic situation. In T.E. Sebeok, ed. *Current trends in linguistics* vol. 8: *Oceania*, 541–657. The Hague: Mouton.
- 1975a, The application of the comparative method to Papuan languages: general and Highlands. In Wurm, ed. 1975:237–261.
- 1975b, Personal pronouns. In Wurm, ed. 1975:191–217.
- 1982, *Papuan languages of Oceania*. *Ars Linguistica* 7. Tubingen: Gunter Narr.
- Wurm, Stephen A., ed., 1975, *New Guinea area languages*, vol. 1: *Papuan languages and the New Guinea linguistic scene*. Canberra: Pacific Linguistics.
- Wurm, Stephen A. and Shiro Hattori, 1981–1983, *Language atlas of the Pacific area*. (2 vols) Canberra, Australian Academy for the Humanities in collaboration with the Japanese Academy.
- Wurm, S.A. and K.A. McElhanon, 1975, Papuan language classification problems. In Wurm, ed. 1975:145–164.
- Wurm, S.A., C.L. Voorhoeve and K. McElhanon, 1975, The Trans New Guinea Phylum in general. In Wurm, ed. 1975:299–322.
- Yen, Douglas E., 1991, Domestication: the lessons from New Guinea. In Pawley, ed. 1991:558–569.
- Z'graggen, John A., 1971, *Classificatory and typological studies in languages of the Madang District*. Canberra: Pacific Linguistics.
- 1975, The Madang-Adelbert Range sub-phylum. In Wurm, ed. 1975:569–612.

- 1980a, *A comparative word list of the Rai Coast languages, Madang Province, Papua New Guinea*. Canberra: Pacific Linguistics.
- 1980b, *A comparative word list of the Northern Adelbert Range languages, Madang Province, Papua New Guinea*. Canberra: Pacific Linguistics.
- 1980c, *A comparative word list of the Mabusu languages, Madang Province, Papua New Guinea*. Canberra: Pacific Linguistics.
- 1980d, *A comparative word list of the Southern Adelbert Range languages, Madang Province, Papua New Guinea*. Canberra: Pacific Linguistics.