Journal of the Linguistic Society of Papua New Guinea
ISSN: 0023-1959

Special Issue 2012

Harald Hammarström & Wilco van den Heuvel (eds.)
History, contact and classification of Papuan languages

Part One

HOW RECONSTRUCTABLE IS PROTO TRANS NEW GUINEA? PROBLEMS, PROGRESS, PROSPECTS

Andrew Pawley
Australian National University
andrew.pawley@anu.edu.au

ABSTRACT
This paper asks how far it is possible to make reliable reconstructions of Proto Trans New Guinea (pTNG), the putative ancestor of more than 400 Papuan languages of New Guinea. In order to make such an assessment we need to answer a number of more specific questions, including: (1) Are there trustworthy criteria for determining membership in TNG? (2) Can a clear-cut determination be made for every language? (3) Can we determine whether a particular reconstructed etymon should be attributed to pTNG or to a later stage? That is to say, is the high order subgrouping of TNG well established? (4) Are the available cognate sets sufficient in number and quality, and widely enough distributed, to arrive at reliable reconstructions of pTNG phonology, lexicon and morphology?

TNG is a valid family in that there is a large set of languages that share features clearly indicating common origin. However, the precise membership and high-order subgrouping of the family remain uncertain and the lexical residues of common origin to which the Comparative Method can be applied are sparse. Furthermore, New Guinea is an area where there has been a great deal of bilingualism, leading to lexical and structural borrowing.

Consideration of such factors led several commentators in the 1970s and 80s to conclude that trustworthy reconstructions of pTNG are unattainable using top-down methods, i.e. by searching for systematic agreements between languages belonging to distantly related subgroups. Contrary to this view, I argue that using a top-down strategy has been essential to achieving initial breakthroughs in reconstructing pTNG. But it is clear that gaining a more complete and finer-grained understanding of the history of the family will require the integration of top-down with bottom-up reconstructive work.

Keywords: Papuan languages, Trans New Guinea, classification, phonological reconstruction, methodology

1. INTRODUCTION

This paper asks how far it is possible to make reliable reconstructions of Proto Trans New Guinea (pTNG), the putative ancestor of more than 400 of the non-

Austronesian ('Papuan') languages of New Guinea.[1] In order to make such an assessment we need to answer a number of more specific questions, including the following:

(1) Is TNG a valid family? Are there trustworthy criteria for determining membership in TNG?
(2) Can a clear-cut determination be made for every language?
(3) Can we determine whether a particular reconstructed etymon should be attributed to pTNG or to a later stage? That is to say, is the high order subgrouping of TNG well established?
(4) Are the available cognate sets sufficient in number and quality, and widely enough distributed, to arrive at reliable reconstructions of pTNG phonology, lexicon and morphology? Or have the lexical and morphological resemblances shared by distantly related TNG languages been so greatly eroded by the passage of time, and so pervaded by undetectable borrowing, as to render reliable reconstruction impossible? That is to say, is the Comparative Method of reconstruction not applicable in this case?

TNG appears to be a valid family insofar as there is a set of languages, call them 'Core TNG', that share features clearly diagnostic of common origin, including a paradigmatic set of pronouns and cognates representing basic vocabulary concepts whose forms, cross-linguistically, show extremely high retention rates (Dyen *et al.* 1967, Holman *et al.* 2009, Pawley 2009a, Tadmor et al. 2010). There are a substantial number of other languages whose claims to membership in TNG are weak or very weak. These are chiefly languages that are of a different structural cast from typical TNG languages, but which show a few resemblances to them in basic lexicon. It is unclear whether such languages are aberrant members of TNG, or distant cousins of the TNG family, or whether they are unrelated to the TNG family but have borrowed from it. The existence of such problematic cases does not impede the task of reconstructing the common ancestor of Core TNG languages.

However, this task is a daunting one. The most distantly related subgroups of Core TNG probably diverged at least 7,000 years ago and possibly several millennia before that, and the residues of common origin to which the Comparative Method can be applied are sparse.[2] The earliest branchings within TNG have not been established and probably will never be. This means that any cognate set with a limited distribution across the 40 or so generally accepted low-order subgroups cannot safely be attributed to pTNG but only to an early stage of TNG (eTNG).

[1] The paper was presented at the conference *History, contact and classification of Papuan languages*, VU University, Amsterdam, 2-3 February 2012. I am indebted to Lyle Campbell, Edgar Suter and Malcolm Ross and, especially, to the editors of this volume, for helpful comments on a draft version and to Jacinta Smallhorn for pointing to additional cognates in the Binanderean languages.
[2] The grounds for this estimate are outlined in Pawley (2005:96-102). See also Denham (2005).

Furthermore, it is generally accepted that New Guinea is an area where linguistic diffusion has always occurred on a large scale, as a result of intermarriage and trade between small language communities (Comrie 1989, Foley 1986, 2010, Ross 1996, Thurston 1987). As a consequence, the argument goes, one can seldom be sure whether apparently cognate words are genuine cognates or loans, or whether apparently cognate grammatical structures are the result of shared inheritance or convergence.

Consideration of such factors led several commentators in the 1970s and 80s to conclude that, if trustworthy reconstructions of pTNG or eTNG are ever to be attained, it will not be done by using top-down methods, i.e. by searching for systematic agreements between languages belonging to distantly related subgroups. The only hope is to build from the bottom up, first aiming for detailed reconstructions of low-order proto-languages, far removed from pTNG, then comparing these in order to reconstruct one step higher, and so on.

Contrary to this view, I believe that using a top-down strategy has been essential to achieving initial breakthroughs in reconstructing pTNG. I will show that such an approach has enabled modest but significant progress to be made in the reconstruction of elements of phonology, lexicon and morphology and in understanding the family tree structure of TNG. However, an exclusively top-down approach provides only a broad-brush picture, leaving many indeterminacies and inaccuracies in the interpretation of the data. I concede that gaining a more complete and finer-grained understanding of the history of the family will require the integration of top-down with bottom-up reconstructive work.

Much of this paper will be about phonological and lexical reconstruction because this is fundamental to the Comparative Method. For reasons of space I will touch only briefly on morphological reconstruction and subgrouping – a detailed review of these domains would require a much longer paper.

At this point a few remarks about the Comparative Method are in order. The Comparative Method is not just a body of methods but underpins a theory of language history. This theory is the genealogical (or family tree or phylogenetic) model of language relationships. The genealogical model is not designed to explain all aspects of the historical development of languages – far from it. But it is the historical linguist's most powerful means for making sense of certain kinds of systematic resemblances between languages. It offers a particular kind of explanation as to why some sets of languages exhibit regular sound correspondences in certain minimal form-meaning units, morphemes and show striking resemblances in morphological paradigms.

The genealogical model rests on several key assumptions.

(a) Languages always change over time.

(b) Languages are typically transmitted through successive generations of native speakers. In this sense, languages show continuity of descent. The idea of continuity of descent does not preclude borrowing from other languages.

Borrowed material simply becomes part of what is transmitted to the next generation.
(c) When sections of a language community become isolated from each other, each isolate becomes a separate descent line, and subsequent changes are passed on only within that community. Successive splits of this kind yield a family tree (a subgrouping or internal classification).
(d) Within a language community, sound changes tend to be regular across the lexicon, i.e. they occur in most morphemes that meet the conditions for the sound change. Related (cognate) morphemes in sister languages can thus be identified and distinguished from chance resemblances by virtue of the fact that they display regular sound correspondences, e.g. word-initial *s* in language A may correspond to *s* in B, *h* in C, *z* in D and zero in E. Given a set of cognate morphemes, all meaning, say, 'blood', it is this regularity of sound correspondences, together with knowledge of the directions that sound changes usually follow, that allows reconstruction of the form for 'blood' in the common ancestor of these languages.
(e) Although there are quite rigorous procedures for determining systematic sound correspondences between sister languages these procedures will not automatically lead to the 'correct' reconstruction of the sound system of the parent language. Often there are a number of competing hypotheses to reckon with. Part of the theory of reconstruction consists of evaluation measures for choosing between competing reconstructions. Choosing which one is most likely to be valid involves appealing to various criteria such as the relative frequencies of particular kinds of sound change, or whether one reconstructed sound system is more natural (conforms better to attested cases) than another, and so on.

An essential part of reconstruction when there are more than two sister languages in a family is the construction of an internal classification or subgrouping representing the sequence in which sister languages diverged. There is no procedure that automatically gives the correct subgrouping or internal classification of a set of three or more related languages. Evidence for subgrouping resides in changes to the proto-language that are shared by a subset of daughter languages. A particular subgrouping hypothesis assumes that the changes took place in a certain sequence. Its claims to be superior to competing hypotheses are measured in terms of accepted evaluation criteria, such as parsimony and the likelihood of a particular kind of shared change occurring independently.

The family tree model of language relationships is an idealisation that at best yields results that approximate actual historical entities and events. It is, for instance, well known that the breakup of a language does not usually take the form of a sharp separation into discrete dialects that develop independently, but is usually gradual, with continuing interaction across a network of dialects, often creating an uneven distribution of innovations.

Genealogical relationship between languages may be posited on other grounds than regularity of sound correspondences. Historical linguists recognise that certain kinds of lexical and morphological resemblances are more likely to be

due to common origin than to chance or borrowing. In particular, certain kinds of morphemes are much more persistent than others, i.e. more resistant to replacement (Dyen et al. 1967, Holman et al. 2009, Pawley 2009, Tadmor et al. 2010). The most durable elements include items of vocabulary that are highly recurrent and represent concepts basic to the human condition, e.g. personal pronouns, certain kinship terms, and terms for certain body parts and bodily functions and processes, and for universal and salient features of the natural environment, such as water, fire, ashes, earth, stone, sun, moon, wood, leaf and lice. Tight-knit morphological paradigms also tend to show some very durable features.

When the case for the TNG hypothesis was first made, at considerable length, in the 1970s, informed commentators remained agnostic (Foley 1986, Haiman 1979, Heeschen 1978, Lang 1976). The fundamental problem, I will suggest, was not that the evidence for TNG was insufficient but that it was not well used. A sketch of the history of the TNG hypothesis and associated research follows.

2 A BRIEF HISTORY OF RESEARCH ON THE TRANS NEW GUINEA

2.1 THE PERIOD 1948-1980

The claim that almost 500 of the more than 700 non-Austronesian languages of New Guinea belong to a single family, now known as Trans New Guinea, was put forward in the 1970s by Stephen Wurm and his associates at the Australian National University (ANU) (Wurm ed. 1975).[3] This ambitious hypothesis had evolved over a period of some 25 years from much more modest beginnings. Most of the Papuan languages, and especially those of the central cordillera of New Guinea remained little documented until after World War II. Arthur Capell (1948-49) set the comparative ball rolling when he pointed to some striking similarities between Kâte and its sister languages of the Huon Peninsula, on the one hand, and some newly documented languages of the nearby central highlands of Papua New Guinea, on the other. In the late 1950s and the 1960s other scholars, chiefly from the ANU research group led by Wurm, posited a remote common ancestry for a number of disparate language groups occupying parts of the central cordillera (Bromley 1967, Wurm 1960, 1964, 1965, 1971, Voegelin and Voegelin 1965, Voorhoeve 1968). The evidence given in support of these claims consisted chiefly of structural resemblances in morphosyntax together with percentages of resemblant words in basic vocabulary. In this period, Alan Healey's comparative

[3] Joseph Greenberg (1971) went further and proposed a vast Indo-Pacific group including all non-Austronesian languages of Melanesia and the Indo-Malaysian Archipelago plus the Southern Andaman Is. group and the languages of Tasmania. I have argued (Pawley 2009) that the case for the Indo-Pacific hypothesis is unconvincing, to say the least. Greenberg correctly identified a number of probable cognates shared by several of his 14 major subgroups of Indo-Pacific, but did not perceive that these several subgroups, all located in New Guinea, form a unit in themselves.

studies of two small groups, Ok (Healy 1964) and Awyu-Dumut (Healey 1970), stand out as pioneering attempts to apply the Comparative Method to TNG languages.

The first study to identify a sizeable and impressionistically convincing collection of putative cognates shared by very distantly related Papuan groups was McElhanon and Voorhoeve (1970). These two scholars compared basic vocabulary in a set of about 70 transparently related languages spoken in the Finisterre Ranges and the Huon Peninsula, in northeast New Guinea, known as the Finisterre-Huon Phylum, and in another clearly related set of about 68 languages, collectively labelled the Central and South New Guinea Stock, located several hundred kilometres away in various parts of central and south-west New Guinea.[4] McElhanon and Voorhoeve pointed to about 90 putative cognate sets shared by the two groups, representing 53 meanings in the Swadesh basic vocabulary list.[5] They also observed that some of these putative cognate sets could be extended to the small Binanderean group of southeast Papua and a small group located around the Rai Coast in Madang Province. It was to these four groups, encompassing around 170 languages, that the name 'Trans New Guinea phylum' was first applied.[6] (See Figure 1, where this proposed grouping is referred to as TNG I.) They also noted that the East New Guinea Highlands group identified by Wurm (1960, 1964, 1965) would probably prove to be part of the Trans New Guinea phylum.

McElhanon and Voorhoeve did not take the further step of trying to establish regular sound correspondences between the putative cognates. Even so, they had made a very strong case for their hypothesis. So many resemblances in basic vocabulary between such geographically distant groups was powerful evidence for their common origin.

A few years later the ANU group published a large volume reporting progress to date in Papuan comparative studies (Wurm ed. 1975). A central chapter in this (Wurm, Voorhoeve and McElhanon 1975) proposed a much expanded version of the 'TNG phylum', encompassing 491 languages, or around 70 percent of all known Papuan languages. (This expanded version is sometimes referred to as TNG III, rather than TNG II, for reasons outlined below.) The group was ranked as a 'macro-phylum' because it united a number of groups that were already ranked,

[4] The name South and Central New Guinea stock was given by Voorhoeve (1968) to a collection of small groups that he argued are genealogically related. There appear to be no good grounds for supposing that these groups collectively form a subgroup of TNG, as was assumed in Wurm (ed. 1975) and Wurm and Hattori (1981-83).

[5] McElhanon and Voorhoeve organise their list under 53 English glosses and this has misled some commentators into saying they cite 53 putative cognate sets. However, in many cases a single English gloss subsumes multiple TNG cognate sets.

[6] The group was called a 'phylum' in accord with the lexicostatistical ranking system and nomenclature adopted by the ANU group, to indicate that the percentage of basic vocabulary shared by the most diverse members was very low. However, no lexicostatistical percentages were computed as part of McElhanon and Voorhoeve's study.

lexicostatistically, as phyla, sharing less than 12 percent of putative cognates in a 200 item basic vocabulary list.

Figure 1: Trans New Guinea I (after McElhanon and Voorhoeve 1970)

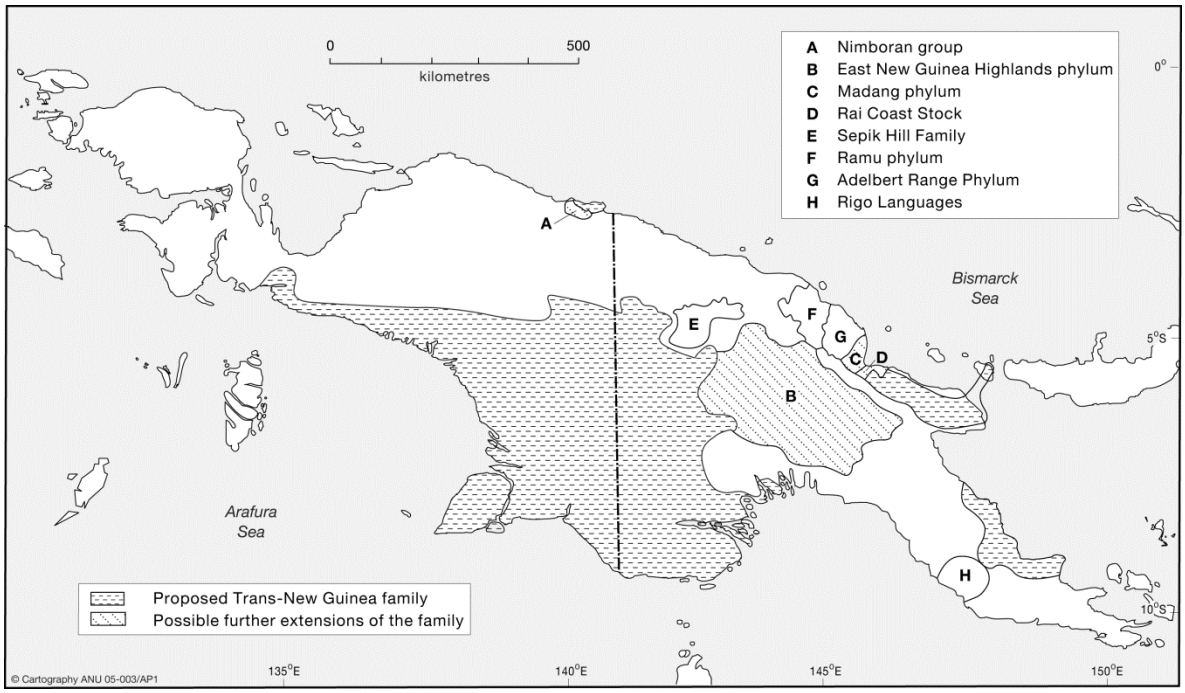
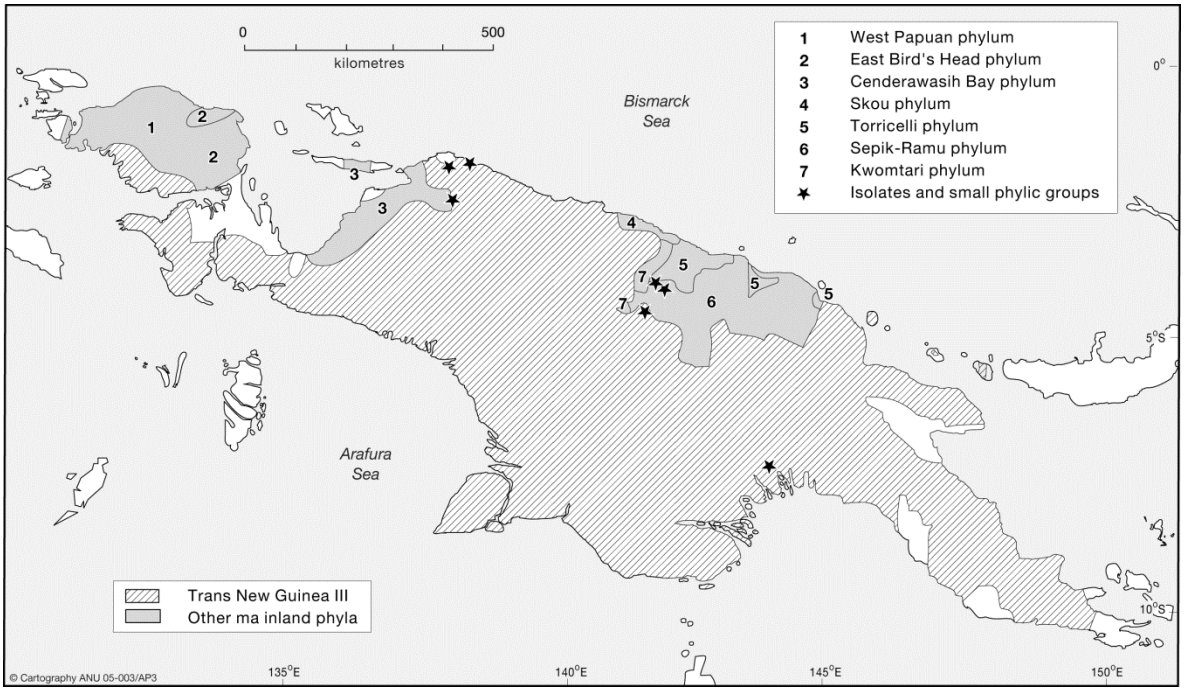


Figure 2: Trans New Guinea III (after Wurm, Voorhoeve and McElhanon 1975)



Why did Papuan specialists remain sceptical of the Trans New Guinea hypothesis?

The 1975 volume received mixed reviews. For the next two decades most specialists in Papuan languages remained agnostic about the validity of all versions of the TNG hypothesis (Haiman 1979, Heeschen 1978, Lang 1976, Foley 1986). Why?

There are several reasons. One is that in the key chapter on the membership of TNG, the authors entered a major qualification. Of the 491 languages they included in TNG, 235, or fewer than half, were viewed as full lineal descendants of Proto TNG. These were referred to as the 'main section' of the macro-phylum. Most of these languages are spoken in the cordillera that runs down the centre of New Guinea from east to west, as far as the Bird's Head. If we label McElhanon and Voorhoeve's grouping as TNG I, this grouping of 235 'main section' languages may be called TNG II.) A further 256 languages were classed as being mixed languages, consisting a TNG component overlaid on a substantial non-TNG substrate. Their membership in TNG was described as "secondary, partial or fractional" (Wurm, Voorhoeve and McElhanon 1975:300).[7] The grounds for this conclusion appear to be either that the semi-TNG languages lack most of the structural features and/or the pronouns that are typical of TNG or that their pronoun paradigms contain some forms that are shared with certain other, non-TNG families (Wurm 1975a). The semi-TNG languages were mainly located outside the central cordillera. In the maps in Wurm and Hattori's (1981-83) linguistic atlas, it is the extended, 491 member version of TNG (TNG III) that is represented.

A further problem was that the claims for a TNG macro-phylum were not underpinned by credible reconstructions of pTNG phonology or lexicon and there was little discussion of shared morphology. The case rested mainly on typological grounds and problematic lexical arguments.

The most valuable evidence for the TNG hypothesis given in the 1975 volume lay, or should have lain, in the pronoun forms. Wurm (1975a) identified three distinct sets of pronouns that are widespread among the Papuan languages. Although no reconstructions were presented for pronoun sets, 1-3, Wurm posited 'basic forms' of each set, which can be regarded as quasi-reconstructions.[8] Set 1 forms are said to be characteristic of full members of TNG and, to a lesser extent,

[7] Wurm, McElhanon and Voorhoeve (1975) refer to the various groups of semi-TNG languages as 'sub-phyla' of TNG, a somewhat confusing choice of term, as it has nothing to do with their lexicostatistical ranking and there was no suggestion that these languages form a subgroup of TNG.

[8] The basic forms of set 1 pronouns were cited as: *na* '1SG', *ka* 2SG', *a* '3SG', *ni* '1PL', *ki* '2PL' (Wurm 1975a, Voorhoeve 1975:449, fn. 32).

of semi-TNG languages. Sets 2 and 3 are chiefly associated with non-TNG families but some forms from these two sets also occur in languages that Wurm regards as semi-TNG.

The pronominal evidence, which should have formed the sharp edge of the case for a Core TNG group, was compromised by its manner of presentation. Although Wurm speaks of the pronoun sets as being characteristic of genetic groupings he muddies the waters by proposing that pronouns were readily borrowed and moved rather freely across language family boundaries (for commentary, see Voorhoeve 1987).

The phonological and lexical evidence, too, was badly handled. Wurm (1975b) presented just nine lexical reconstructions attributed to pTNG, each based on widely distributed sets of putative cognates, and proposed a set of pTNG consonant and vowel phonemes. Unfortunately, these reconstructions do not bear close scrutiny. No tables or examples showing sound correspondences between pTNG and particular daughter languages are provided. A plethora of consonant and vowel phonemes, represented by idiosyncratically chosen symbols, is reconstructed without any indication as to how these might form a plausible consonant or vowel system. Several of the reconstructed lexical forms are formulas that contain so many variable phonological segments as to be both highly unrealistic and unfalsifiable, e.g. the reconstruction for 'eye' is **(avu-)Da$^{>\,<}$K^{M}A(Pur)$^{(n)}$*, where $^{>\,<}$ indicates metathesis of the adjacent syllables.

It seems that all these variable elements in lexical reconstructions were intended to allow a diverse array of resemblant forms to be derived as regularly corresponding reflexes, without appeal to natural phonological and morphological processes that might have yielded irregular changes. Wurm does not list the putative reflexes of this reconstruction for 'eye' but he does so for some other reconstructions. I will comment on just a couple of typical examples.

Languages of diverse subgroups from all quarters of the TNG area have a verb root *na-* 'eat'. It would seem a straightforward matter to attribute this form to pTNG. However, instead of **na-* Wurm reconstructs **(Ï)(n,N)a$^{(d)}$A(i)*. This is evidently done in order to allow forms such as *navai*, *una, inda, da, za, ane, nai*, *laia* and *idie*, to be regarded as reflexes of the same pTNG etymon. It would make better sense to assume the pTNG root was **na-* and that (a) some of these other forms contain the root **na-* fused with other morphological material and (b) some are not cognate.

Similarly, there is a pTNG etymon we can reconstruct as **inja* or **ita* 'tree, wood, fire' (where *t may be [r] intervocally), which languages from diverse groups reflect variously as *ita, ija, ira, ila, ida, inda, ita, dza, izi, dzi.* Wurm prefers to reconstruct **(Ï)DADe* or **(Ï)DAkaP$^{(M)}$A* (where A and D carry diacritics I have omitted), evidently in order to be able to derive such diverse forms such as *irama, iduk, dika, sare, tedzi, tombe, kera* and *tokwa*, some of which are probably not cognate and others of which may represent a fusion of two morphemes.

These reconstructions were arrived at using a top-down approach. As already noted, reviewers concluded that it was not possible to make useful reconstructions of pTNG using top-down methods. Underlying this conclusion was the belief that widespread cognate sets were too few, relationships between subgroups too remote and uncertain, and borrowing too pervasive, to allow regular sound correspondences or other systematic resemblances to be established.

2.2 THE PERIOD SINCE 1980

Little new historical work was done on TNG during the 1980s. William Foley (1986) devotes a lengthy chapter to Papuan comparative-historical studies in general but his main contribution to Trans New Guinea historical linguistics in this is an examination of the lower order groups he calls Gorokan and Kainantu. In this chapter Foley expressed serious reservations about the value of the family tree model for doing Papuan historical linguistics.

> 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 to the smooth application of the family tree model 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. (Foley 1986:209)

Foley notes that historical linguists, when faced with distinguishing between borrowings from true cognates, generally rely on the sound correspondences found in basic vocabulary: "pronouns, nouns referring to body parts, simple kin relations, natural phenomena like the sun, moon, stars, rain, trees, fire, water, mountains; and verbs of bodily actions, like eating, hitting and giving" (1986:211). The trouble is, he asserts, in Papuan languages even basic vocabulary is prone to borrowing. No noun or pronoun is safe (he allows that basic verb roots are probably more resistant to borrowing than nouns). Thus the usual tests for distinguishing borrowed from inherited words cannot be relied on.

Foley was right to emphasize the challenge that borrowing presents in applying the family tree model to Papuan languages but his generalisations about the extent of borrowing in basic lexicon were probably too sweeping. There are undoubtedly cases of extensive borrowing in basic vocabulary among Trans New Guinea languages (e.g. Comrie 1986, 1989, Shaw 1986) but the examples that he cites of borrowing in pronouns and other basic vocabulary have since been discredited (Chowning 1987, Ross 2005).[9] As far as the reconstructability of pTNG is concerned, the proof of the pudding must lie in the tasting.

[9] Chowning (1987) offers an incisive critique of claims that a number of widespread TNG cognate sets reflect ancient borrowings from Austronesian sources before the dispersal of TNG. Ross

In the mid 1990s a small group of ANU scholars, chiefly Malcolm Ross, Meredith Osmond, Edgar Suter and myself, returned to the challenge of reconstructing elements of pTNG or eTNG. A large part of pTNG segmental phonology was reconstructed, along with a body of lexical items which now stands at around 200, mainly using a top-down approach. The evidence for the phonological reconstructions will be discussed in detail in sections 4 and 5.

Building on earlier work by McElhanon and Voorhoeve (1970), and Wurm (1975b), Ross reconstructed free pronoun paradigms both for pTNG and for the proto-languages of many subgroups of TNG. The free pronoun forms Ross (2005) attributes to pTNG are given below. Note the pattern whereby the consonant remains constant in the corresponding persons with the singular plural contrast marked by a vowel difference, typically *a (singular) vs. *i (plural).[10]

Table 1 pTNG free pronouns

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

Ross (1995, 2000, 2005) uses personal pronoun forms as a preliminary diagnostic for determining TNG membership.[11] Although pronouns had played an important part in Wurm's account of the history of the Papuan languages, Ross's approach differs from Wurm's in certain fundamental ways. Whereas Wurm held that pronouns and even whole pronominal paradigms are readily borrowed, Ross argues forcefully that systematic resemblances in pronominal paradigms are a particularly powerful form of evidence for genetic relationship and for subgrouping purposes because such paradigms are known to be highly resistant to borrowing.[12] Ross argues that if a Papuan language reflects two or more of the pTNG personal pronoun roots, this is strong evidence that it belongs to TNG. Among the 605 Papuan languages he surveyed some 311 met this criterion (Ross

(2005:53-58) points out that claims concerning extensive borrowing of pronouns among Papuan languages are not well founded.

[10] The existence of a canonical set of TNG pronouns had already been noted by McElhanon and Voorhoeve (1970) and Wurm (1975).

[11] On the basis of the pronominal evidence Ross (2000, 2005) identifies some 23 Papuan language families which appear to be unrelated, TNG being just one of these.

[12] A drawback to using TNG pronoun forms as evidence of common origin or subgrouping is that they are short, typically having the form CV, raising the possibility that formal resemblances are due to chance or to universal tendencies favouring the use of certain consonants in pronouns. However, the larger the number of pronominal agreements the less likely it is that these factors can explain the agreements. See Rhodes (1997) and Ross (2005) for further discussion.

2000). Another 36 did not satisfy it but were assigned to TNG because their basic vocabulary includes a number of well-established TNG lexical forms. Another 12 languages were seen as having even more marginal claims to membership in TNG. When daughter languages show changes to the reconstructed pTNG pronominal paradigm, Ross interprets this as evidence for defining subgroups rather than as evidence for substrate effects.

Among the cognate sets of free pronouns there is also evidence for pTNG dual pronouns. Languages in several widely scattered subgroups have dual pronouns that reflect a dual suffix. And among languages that lack a dual/plural contrast there are some whose plural pronouns that appear to reflect old dual forms. Following Haiman (1979), Ross (2000:77, 158-160) reconstructs the dual marker as *-li, with an alternant *-t, although he does not say whether this marker could be combined with all the singular pronominal roots. I reconstruct *-Li, *L representing a consonant that may have been *t, *l or a lateral that was distinct from *l. The full range of evidence suggests that there were free form dual pronouns having the approximate forms: *niLi or *nuLi '1st plural', *ŋgiLi or *kiLi '2nd dual' and *iLi '3rd dual'.

Edgar Suter (t.a.) has shown that a set of object pronouns serving as clitics or prefixes to transitive verbs can be reconstructed for pTNG and that in some languages, the object pronoun clitics have been better preserved than the free form pronouns. Suter (1997) and Pawley (1966, 2000) have reconstructed fragments of early TNG verb morphology. There is support from this quarter for a dual/plural contrast in the free form pronouns: such a contrast is marked in the verb desinences which mark person-and-number of the subject of the verb. Most of the pTNG verbal suffixes appear ultimately to be cognate with the pTNG free form pronouns.

3. SUBGROUPING ISSUES

By the 1970s upwards of 40 more or less transparent subgroups of TNG languages had been identified, most with just a few member languages. Only two of these groups contain more than 20 members: the Finisterre-Huon group, centred in Morobe Province but extending into eastern Madang Province, comprises about 70 languages, and the Madang group, spoken almost exclusively in Madang Province, comprises about 100 languages. These two large groups both have a number of diverse branches. The internal lexicostatistical and structural diversity of the Finisterre-Huon and Madang groups is such as to suggest that their respective proto-languages broke up at least four millennia ago and, in the case of the Madang group, perhaps well before that.

How the generally accepted subgroups relate to one another at higher levels remains a matter of debate. Ross (2000, 2005) has posited a number of higher-order groups within TNG based on shared innovations in free pronoun paradigms. Voorhoeve (2005) has tentatively proposed that Ok and Awyu-Dumut form a subgroup as opposed to Asmat-Kamoro, on the basis of much closer morpho-

syntactic similarities. However, at this stage none of these putative higher-order groupings can be considered well-defined.

The following is a list of TNG subgroups that are (a) generally accepted or (b) remain controversial but are included because they show reflexes of two or more pTNG personal pronouns and/or reflexes of ten or more other basic vocabulary items .[13] Most subgroups are listed under geographical regions, (alphabetically ordered) for convenience of locating them on the relevant map (Figure 3). The exceptions are the Angan, Finisterre-Huon and Madang groups, which stand alone. There is no implication that any of the regional collections corresponds to a subgroup.

The ten languages whose phonological histories are examined in section 4 are selected from among the generally accepted subgroups.

TNG subgroups listed by region

ANGAN GROUP (Gulf, Eastern Highlands and Morobe Provinces, PNG)

SOUTH BIRD'S HEAD
Inanwatan-Duriankere, Konda-Yahadian, South Bird's Head Proper: Arandai, Barau, Kampong Baru, Kasuweri, Puragi, Tarof

BOMBERAI PENINSULA (Papua, Indonesia)
Tanah Merah, West Bomberai

CENTRAL HIGHLANDS OF PAPUA NEW GUINEA (Eastern Highlands, Enga, Simbu, Southern Highlands, Western Highlands Provinces)
Chimbu–Wahgi group, Duna-Bogaia group, Enga-Huli group, Kenati
Kainantu-Goroka subgroup: Kainantu subgroup, Gorokan subgroup, Wiru

FINISTERRE–HUON GROUP (Morobe and eastern Madang Provinces, PNG)
Finisterre group: Erap subgroup, Gusap–Mot subgroup, Uruwa subgroup, Wantoat subgroup, Warup subgroup, Yupna subgroup
Huon Peninsula group: Eastern Huon subgroup, Western Huon subgroup

MADANG GROUP (Madang Province, PNG)
Croisilles Linkage (includes most of Z'graggen's (1975, 1980) North Adelbert Range group plus his Mabuso group), Kalam-Kobon subgroup, Rai Coast subgroup, South Adelbert Range subgroup

SOUTH-CENTRAL NEW GUINEA (Western and Gulf Provinces of PNG and southeast margin of Papua, Indonesia)
Gogodala-Suki group, Inland Gulf group, Kiwaian group, Moraori, Tirio, Turama-Kikori group

[13] In identifying lower order subgroups here, I largely follow Wurm and Hattori (1981–83). However, I have in some cases used different names for subgroups, where there seem to be good reasons to do so. For a more elaborate discussion on TNG subrouping see Pawley and Hammarström (t.a.).

SOUTH-EAST NEW GUINEA (Central, Milne Bay and Northern Provinces of PNG)
Binanderean-Guhu Samane (or Greater Binanderean) group: Binanderean subgroup and Guhu–Samane, Dagan group, Goilalan group, Koiarian group, Kwalean group, Mailuan group, Manubaran group, Yareban group

SOUTH-WEST NEW GUINEA (Papua Province, Indonesia, south of the central highlands, roughly from the Digul shelf to the neck of the Bird's Head)
Asmat–Kamoro group, Awyu–Dumut group, Kayagar group, Kolopom group, Marind group (Boazi, Marind and Yaqay subgroups), Mombum group

NORTH-CENTRAL NEW GUINEA (northeast Papua Province, Indonesia, just west of Jayapura)
No generally groups

WESTERN HIGHLANDS OF PAPUA NEW GUINEA (chiefly, the highlands of PNG just east and west of the Strickland Gorge, but extending as far west as the eastern margins of Indonesian Papua)
Awin–Pa group, Bosavi (Bedamini) group, East Strickland group, Kamula, East Kutubu group, West Kutubu group, Ok-Oksapmin group: Ok subgroup, Oksapmin

WEST PAPUAN HIGHLANDS (highlands of Papua Province, Indonesia, excluding the Ok group)
Dani group, Mek (= Goliath) group, Uhunduni, Wissel Lakes group

In addition there are a number of other groups and isolates whose classification remains controversial. These include

Groups whose TNG status is uncertain

BOMBERAI PENINSULA
Mairasi group, Mor

CENTRAL HIGHLANDS OF PAPUA NEW GUINEA
Pawaian

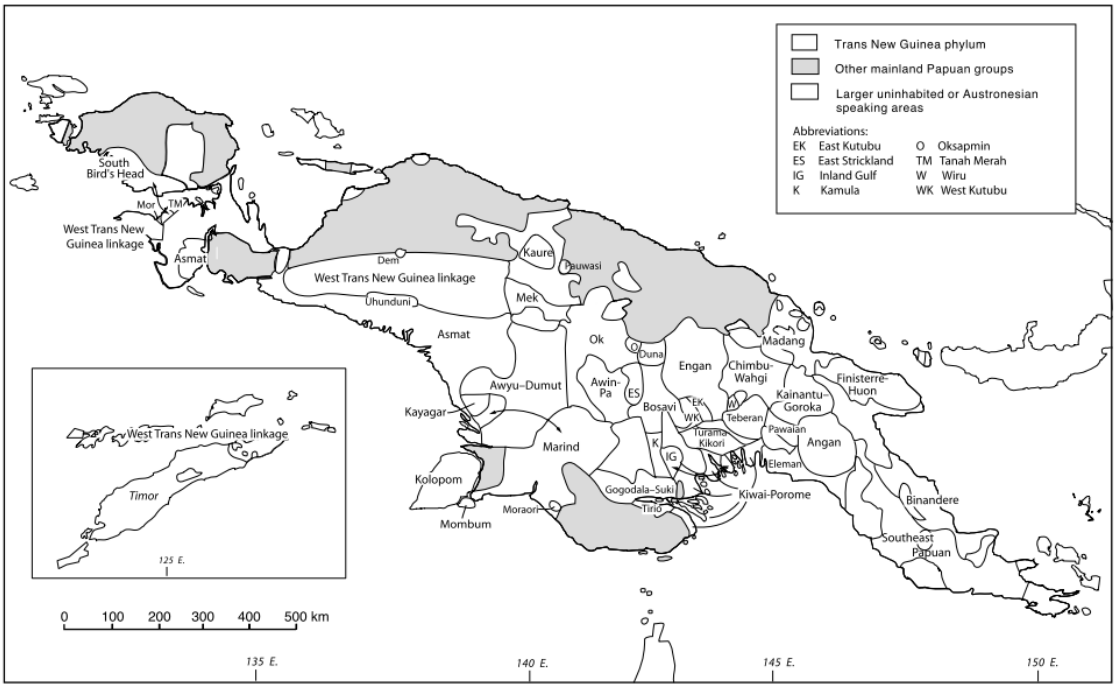
NORTH-CENTRAL NEW GUINEA
Karkar-Yuri, Kaure group, Pauwasi group, Sentani group

SOUTH-CENTRAL NEW GUINEA
Porome, Teberan group

TIMOR REGION
Alor-Pantar, West Timor, East Timor

WEST PAPUAN HIGHLANDS
Dem

Figure 3: Locations of putative subgroups of TNG (after Ross 2005)



4. RECONSTRUCTING pTNG SEGMENTAL PHONOLOGY

4.1 PROCEDURES

In ideal circumstances, given a large team of researchers, ample data and a well-defined family tree to work with, one would employ a combination of bottom-up and top-down strategies to reconstruct pTNG phonology and lexicon. Using a bottom-up approach one would first compare members of each well-defined low-order subgroup and reconstruct the proto-language or interstage ancestral to these groups, and then proceed upwards stage by stage through intermediate-level subgroups until one reaches pTNG itself.

However, in the present circumstances, there are several practical objections to using a bottom-up strategy to reconstruct pTNG phonology. First, for most low-order subgroups the paucity of good descriptions of member languages means that detailed bottom-up reconstructive work is simply not possible at present, and probably will not be for decades to come. Second, with a few exceptions, the proto-languages of most of the generally accepted low-order subgroups have a shallow time depth. They are so far removed from pTNG that reconstructing them will not tell us much about pTNG that cannot be inferred from comparisons of the living languages.[14] A further serious problem is that, as we move up the tree, putative higher-order subgroups are (with a few exceptions) not well-defined.

[14] Possible exceptions are the Finisterre-Huon, Madang and Ok groups, each of which shows a degree of lexicostatistical diversity consistent with a time depth of five millennia or more.

For these reasons, when I began trying to reconstruct pTNG phonology I chose for the most part to follow a top-down approach, comparing a sample of contemporary TNG languages drawn from diverse lower-order subgroups.

The top-down method, applied to the TNG family, has major limitations. It yields reconstructions that are approximate, broad brush, leaving some details indeterminate. For another thing, the pool of widely distributed cognate sets, and therefore the pool of reconstructions that can safely be attributed to pTNG, is small. The first step in applying the Comparative Method is usually to search for regular sound correspondences between pairs of attested languages. However, a comparison of any pair of contemporary TNG languages drawn from two of the 50 or so generally accepted subgroups is likely to yield between 10 and 20 likely cognates – too few to discover many recurrent sound correspondences. I therefore adopted a strategy designed to yield a larger body of cognate sets that could be searched for sound correspondences. This involved the following operations.[15]

1. A pool of putative cognate sets was assembled based on comparisons of many disparate subgroups. McElhanon and Voorhoeve (1970) had built a valuable platform when they compared basic vocabulary in Finisterre-Huon languages with that in half a dozen subgroups located in central New Guinea, later including data from the Binandere and Rai Coast groups. Of the 90 possible cognate sets they proposed, most were convincing. The publication after 1970 of word lists for many more languages made it possible to extend this base. Of particular importance was Z'graggen's (1980a-d) extensive set of wordlists for Madang languages.

2. The putative cognate sets were searched for sound correspondences that spanned languages from different subgroups.

3. A sample of languages was chosen for closer inspection. These were languages that (a) participated in a relatively large number of widespread cognate sets and (b) represented a range of geographically dispersed subgroups. The choice was not a straightforward matter because languages that best fitted criterion (a) were mainly located in the eastern half of New Guinea. The strongest looking (most recurrent) correspondence sets found in this sample of languages were used to reconstruct a set of pTNG segmental phonemes, which were assigned approximate phonetic values. Other correspondence sets were placed aside, as residue to be explained later.

4. This putative phonological system was compared with phonological systems of contemporary TNG languages to assess its plausibility as a structural type and certain correspondence sets were re-evaluated accordingly. For example,

[15] This attempt was due to an invitation received in December 1994 to contribute to a festschrift for Bert Voorhoeve. (Baak et al. 1995) Wracking my brains for a suitable topic, it occurred to me that the cognate sets in McElhanon and Voorhoeve (1970) provided a good place to begin reconstructing the segmental phonology of pTNG, something that until that time I had given little thought to.

although the glides *w* and *y* figured in very few widespread cognate sets, and in that respect were poorly supported, these glides are so widely distributed among contemporary TNG languages that it would surprising if they did not have counterparts in pTNG.

5. Approximate reconstructions were made of full lexical forms in pTNG and eTNG, based both on the proposed segmental phonemes and on hypotheses about pTNG syllable and morpheme structure.

6. The next step was to try to establish regular correspondences between pTNG phonemes and their reflexes in the daughter languages. Comparing daughter languages with pTNG, rather than with each other, enlarged the number of comparisons yielding sound correspondences.

7. As work proceeded, all reconstructions were reviewed in the light of additional observations. Reconstructed forms were revised until they best fitted their putative reflexes in daughter languages.

Findings were reported in Pawley (1995, 2001, 2005, 2011). Unsurprisingly, the results yielded by this procedure are far inferior to those achieved in the study of language families such as Indo-European and Austronesian, where the pool of cognate sets is much larger and which have been intensively researched. In many cases the forms of reconstructed etyma show multiple indeterminacies because the reflexes in daughter languages do not agree. Nevertheless, I think the results represent an advance in our understanding of TNG historical phonology.

Before reading the rest of section 4, the reader may wish to skip ahead to section 5, which contains an account of how several lexical reconstructions were arrived at.

4.2. OUTLINE OF pTNG SEGMENTAL PHONOLOGY

4.2.1 CONSONANTS AND VOWELS: AN OVERVIEW

Table 2 shows the consonant and vowel phonemes so far attributed to pTNG, based on correspondences among representatives of several diverse subgroups.

Table 2 pTNG segmental phonemes

consonants

	labial	apical	laminal	velar
oral obstruents	p	t	s	k
prenasalised obstruents	mb	nd	nj	ŋg
nasals	m	n		ŋ
lateral		l		
glides	w		y	

vowels

	front	central	back
high	i		u
mid	e		o
low		a	

The symbols for particular pTNG phonemes should not be taken as a claim that these phonemes had constant phonetic values. Phonemes typically have allophones that are either context-dependent or in free variation. For example, many TNG languages show 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] or [p, ɸ β]. Many languages have no contrast between the apical stop [t] and flapped [ř].

There is no suggestion that the 14 consonants and five vowels in Table 2 are an exhaustive list of pTNG segmental phonemes. They represent a first approximation: a phonological inventory based on more or less well attested correspondence sets. There is some evidence for additional phonemes, for example, for a contrast between *t and *r and for a second lateral phoneme *L. *L is used below in certain reconstructions where some correspondences point to *t and others to *l.

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. Some of the problematic correspondences are can be explained as due to natural processes of change (assimilation, dissimilation, loss of unstressed vowels or syllables, analogical reshaping, etc.) where these have applied not regularly but sporadically, affecting only a few of the eligible forms. It is worth noting that certain kinds of sporadic sound changes, which might be called 'one-step feature shifts', are common in TNG languages. In such cases the pronunciation of a segment shifts from one phoneme to another, differing from the first only in one distinctive feature. Furthermore, only certain pairs of features may be related in this way. For example, the following sorts of shifts recur:

Table 3: Some recurrent irregular sound shifts in TNG languages

apical nasal to velar	lateral to nasal	nasal to lateral
n > *ŋ*	*l* > *n*	*n* > *l*

glide to fricative/affricate	apical stop to sibilant and vice versa
y > *z, dz*	*t* > *s, s* > *t*

front to back	back to front	high to mid	mid to high
i > *u*	*u* > *i*	*i* > *e*	*e* > *i*
e > *o*	*o* > *e*	*u* > *o*	*o* > *u*

TNG languages are in no way unusual, among language families, in exhibiting a plethora of irregular correspondences. Greenberg reminds us that that the same thing is found in Indo-European. He lists 82 cognate sets that French and English have independently inherited from pIE and specifies all the initial and final consonant correspondences and all the vowel correspondences (Greenberg 2006:91-3). He finds that

> An examination of the initial consonants, the most regular of all, shows for example that *k* in French corresponds to English *f* (once), *h* (three times) *hw* (twice) and *j* (once); and so for all the other initial consonants to a lesser or greater degree. The vowels are close to statistical randomness, and in this they are rivalled by the final consonants. In the light of the earlier discussion the reasons should be obvious. As time goes by these accumulate, along with various sporadic phonetic processes, grammatically induced analogy, and new derivational processes, while lexical replacement reduces the number of cognates. All of these processes work to produce more and more diversity of correspondence and virtually never to reduce it. (Greenberg 2006:92-3)

Greenberg is not saying that the principle of regularity of sound correspondence is useless for identifying cognates shared by English and French, or any pair of languages that separated many millennia ago. His point is simply that, when time depth is great, many valid cognates will fail to show regular correspondences. It follows that instead of discarding resemblant forms that do not correspond regularly, as due to borrowing or accident, we should search for possible explanations of the irregularities in terms of language-internal processes. The problem is that, when knowledge of the history of the languages in question is very sketchy, it is often hard to arrive at plausible explanations or to evaluate such explanations.

Notational conventions

The following notational conventions are used here in representing reconstructed forms:

C consonant of indeterminate form
V vowel of indeterminate form
(x,y) a segment is reconstructed but is indeterminate between x and y
[x] it is uncertain whether a segment should be reconstructed in this position
*tumuk/*kumut alternative reconstructions. It is uncertain which should be preferred

The following convention is used in reflexes of reconstructed morphemes:

niman(amp) (amp) is non-cognate material, not part of the reflex

4.2.2 THE STRUCTURE OF SYYLABLES AND LEXICAL ROOTS

Syllables had the shape (C)V and in word-final position they could also be (C)VC. The prenasalised obstruents evidently did not occur syllable-finally. There were probably no phonemic consonant clusters within words, homorganic nasal + obstruent clusters, such as [mb], being interpreted as unit phonemes. It is likely that vowel clusters were not permitted, as opposed to sequences of glide + vowel or vowel + glide.

Lexical roots could consist of one or more syllables. Verb roots often consisted of a single syllable, e.g. **na-* 'eat, drink', **nVŋg-* 'know, hear, see', **tV-* 'do, make'. A few nouns consisted of a single syllable, e.g. **m(o,u)k* 'milk, sap, breast', **mV* 'taro'. Singular and plural pronouns were also monosyllabic (see Table 1).

A high proportion of reconstructed roots consist of two syllables, either with final syllable open: **mbena* 'arm', **imbi* 'name', **ke(nj,s)a* 'blood', or closed: **iman* 'louse', **kend(o,u)p* 'fire', **mbalaŋ* 'flame', **mundun* 'internal organs'.

A fair number of roots consist of three syllables, e.g. **kumbutu* 'wind', **kutV(p,mb)(a,u)[C]* 'long', **(ŋg,k)andapu* 'skin, bark', **mun(a,i,u)ka* 'egg', **(ŋg,k)atata* 'dry', and some perhaps of four, e.g. ** tututu[ŋga]* 'straight'.

In some cases it is uncertain whether a pTNG root was disyllabic or trisyllabic. Although some daughter languages have a trisyllabic reflex, the final vowel may have been added as an echo-vowel to reinforce a preference for final open syllables. Reconstructed elements which may or may not have been present in the proto-form are enclosed in square brackets, e.g. **takVn[V]* 'moon', **maŋgat[a]* 'teeth, mouth'.

4.2.3 STRESS PLACEMENT IN POLYSYLLABIC ROOTS

It will be suggested in 5.3.3 that pTNG may have had contrastive placement of stress. The argument rests on the variable outcomes of pTNG disyllabic and trisyllabic roots in certain daughter languages.

4.3 REFLEXES OF pTNG PHONOLOGICAL RECONSTRUCTIONS IN A SAMPLE OF TEN LANGUAGES

4.3.1 The sample languages

In the following sections attested reflexes of the segmental phonemes attributed to pTNG are given for a sample of ten languages, together with supporting cognates. The languages, together with their subgroups are as follows.[16]

[16] Sources of lexical data for the sample languages are:

Language	Sources
Apali (=Emerum)	Wade n.d., Z'graggen 1980d
Asmat	McElhanon & Voorhoeve 1970, Voorhoeve 1965

Middle Wahgi (Chimbu-Wahgi group)

Kalam (Kalam-Kobon branch of the Madang group.) Kalam has two major dialects: Etp and Ti. Data cited here are from Etp, unless otherwise indicated.

Apalɨ (aka Emerum) (Sogeram subgroup of S Adelbert Range branch of the Madang group). Apalɨ has two sharply distinct dialects, Acɨ and Akɨ, which differ from each other in their reflexes of several pTNG phonemes. Wade (n.d.) gives forms for both dialects whereas the forms in Z'graggen (1980d) are evidently from Acɨ alone. Data cited here are mainly from Akɨ.

Selepet (W. Huon branch of the Finisterre-Huon group)

Kâte (E. Huon branch of the Finisterre-Huon group). Kâte has several dialects. Data cited here are from the Wemo dialect.

Binandere (Binanderean group)

Telefol (Mountain Ok subgroup of the Ok group)

Kiwai (Kiwaian group). Kiwai has several dialects. Data cited here are from Island Kiwai unless otherwise indicated.

Kaeti (Dumut branch of the Awyu-Dumut group)

Asmat (Asmat branch of the Asmat-Kamoro group). The name Asmat is applied to several dialects or closely related languages. Data here are cited from Flamingo Bay Asmat unless otherwise noted.

The location of these languages is shown in fig. 2. I would have liked to have included more languages from the West Papua, such as Mek, Grand Valley Dani and Ekagi, but the numbers of eTNG reflexes detected in the limited data I have inspected for these languages were too small to have a chance of establishing regular sound correspondences.

Almost 200 etyma have been attributed to an early stage of TNG (eTNG) because they are reflected in two or more subgroups that are not known to belong to a higher-order subgroup of TNG (Pawley 2011). It is to be expected that some

Binandere	King 1927, Wilson 1969, McElhanon & Voorhoeve 1970, Smallhorn 2011
Kaeti	McElhanon & Voorhoeve 1970
Kalam	Pawley and Bulmer (2011)
Kiwai	Franklin 1973 (Appendix), McElhanon & Voorhoeve 1970, Wurm 1973
Kâte	Flierl and Strauss 1977, McElhanon 1967, McElhanon & Voorhoeve 1970
Middle Wahgi	Ramsey 1975
Selepet	McElhanon 1967, McElhanon and McElhanon 1970, McElhanon & Voorhoeve 1970
Telefol	Healey 1964, Healey and Healey 1977, McElhanon & Voorhoeve 1970

of these etyma will turn out not to be valid, because the putative cognate sets which they are based on involve chance resemblances or borrowing.

Of the 188 eTNG reconstructions examined, 100 are found in both the eastern and western halves of New Guinea. For present purposes the dividing line between eastern and western New Guinea approximates the border between Papua New Guinea and West Papua and groups that straddle the political border, such as the Ok, Marind and Pauwasi groups, are assigned to the western half.

The geographic distribution of the remaining lexical reconstructions shows a strong bias. Eighty-four reconstructions have reflexes in subgroups confined to the eastern half of New Guinea. Just four reconstructions are attested only in western New Guinea. This bias probably reflects two factors: (i) there are considerably more TNG languages in the eastern half of New Guinea than in the western half, (ii) I have searched more diligently among eastern languages than among western languages.

Although close to 200 lexical reconstructions are attributable to an early stage of TNG, no one language today retains more than a small proportion of these. The largest number of reflexes of eTNG etyma noted for any one language is around 40, for Kalam. In some putative TNG languages for which data are scanty, it is difficult to find more than four or five reflexes. Even in the case of languages with good dictionaries one can often find only 20 to 30. Such a paucity of cognates is one of the things that led Wurm to conclude that many TNG languages have only a veneer of TNG laid upon an unrelated substrate language.

Now 20 reflexes, or even 100, 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 comparisons come into play. In some cases a contemporary language belongs to a lower-order subgroup within which it has close relatives. Between them the members of a sizeable subgroup will have more reflexes than any single language in the group. In such cases, it is sometimes possible to use the sound correspondences exhibited by members of the subgroup to extend the range of correspondences between pTNG and any one contemporary language.

In the presentation of sound correspondences that follows, it will be seen that in many cases a particular pTNG phoneme has multiple reflexes in a certain language, e.g. *k may be reflected as *k* in some words, as *g* in others and as zero in others. It is not claimed that each reflex is a *regular* correspondence. The main purpose of the present exercise is to present the data in an orderly manner. Explaining apparently irregular correspondences is a task for the future. When working with just a few putative cognate sets, we cannot afford to discard promising comparisons just because they show unexplained irregularities. It is acknowledged that this approach has its dangers – irregular spurious comparisons will sometimes be included.

A number of abbreviatory conventions are used in representing reconstructions.

[x]	it cannot be determined whether x was present
(x,y)	either x or y was present
C	an indeterminate consonant
V	an indeterminate vowel
xyz-	xyz is a bound form, normally taking a suffix

4.3.2 Evidence for the consonants

4.3.2.1 Nasals

Two of the three nasals, *m and *n, are well attested, being regularly reflected in many cognate sets in both word initial and medial position. These are probably the two most stable consonant phonemes, in terms of continuity of phonetic character and regularity of reflexes. The third nasal, *ŋ, looks fairly secure even though there are only a few good cognate sets supporting it. However, it is likely that *ŋ was very rare word-initially, occurring only in onomatopoeic and nursery words.

Table 4 shows the frequencies of nasal consonants in eTNG etyma in word initial, medial and final positions.

Table 4 Frequencies of nasal consonants in eTNG lexical reconcontructions

	*m	*n	*ŋ
Initial	23	10	1
Medial	13	17	2
Final	6-8	7	5

It is noteworthy that there are 23 eTNG reconstructions with initial *m, 10 with initial *n and one with initial *ŋ, a pattern in line with predictions made on the basis of frequencies in a sample of contemporary languages. It is also noteworthy that the relative frequencies are different for medial and final positions. For example, *ŋ is more frequent word-finally than word-initially, while for *m the reverse holds.

Table 5 summarises the most regular (or in some cases, the sole instances of) reflexes of pTNG nasal consonants in the languages in the sample. Reflexes are given separately for word initial, medial and final position, where known.

Table 5 Reflexes of pTNG nasals

pTNG m-	**-m-**	**-m**	**n-**	**-n-**	**-n**	**ŋ-**	**-ŋ-**	**-ŋ**
M. Wahgi	m	m	m	n	n	n	ŋ	ŋ
Kalam	m	m	n	n	n	ŋ	–	ŋ
Apali	m	m	–	n	n	n	–	ŋ
Selepet	m	–	n	n	n	ŋ	–	m, ŋ
Kâte	m	m	–	n	n	ŋ	–	–
Binandere	m	m	?	n	–	Ø	–	–
Kiwai	m	–	n	n	–	–	–	–
Telefol	m	m	n	n	–	–	–	ŋ
Kaeti	m	m	–	n	n	–	–	–

It can be seen that these languages have consistently preserved *m and *n in initial and medial positions. Those languages that keep final reflexes also retain *m and *n in final position (Kâte is an exception, having merged all final nasals to *ŋ). *ŋ has also been preserved in the few instances where it is attested.

The nasal correspondences set out here offer no evidence for high-order subgrouping. However, the behaviour of the nasals in particular etyma does offer some such evidence, to be touched on in section 5.

There follows fuller information about reflexes of the pTNG nasals in each of the sample languages.

Middle Wahgi

Middle Wahgi has three nasals, *m*, *n* and *ŋ*. pTNG *m is regularly continued as *m*, at least in word-initial and medial positions.

initial *m > m

*ma- ‘NEG clitic’	ma ‘no!’
*ma(n,k,L)[a] ‘ground’	mał
*maŋgV ‘compact round object’	muŋ ‘fruit, nut, lump’ (cf. also muŋgum ‘kidney’)
*mo(k,ŋg)Vm ‘joint’	mokum, mokem ‘knuckle, joint’
*mundun-maŋgV ‘heart’	mund-muŋ
*mV ‘taro’	mi
*mV(k,ŋ)V[C] + t(e,i)- ‘vomit’	mek (si-) ‘vomit’, mek ‘vomitus’

medial *m > m

*am(a,i) ‘mother’	ama
*amu ‘breast’	am
*niman ‘louse’	numan
*n(o,u)man ‘mind, soul’	numan ‘thought, mind, will’

medial *m > mb

*kumV- ‘die’	kumb- ‘(of fire) die’

final *m > m

*mo(k,ŋg)Vm ‘joint’	mokum, mokem, (angeł) mokem ‘knuckle, joint’

pTNG *n remains *n* in Middle Wahgi initially and finally. No reflexes of medial *n have been found.

initial *n > n

*na- ‘eat’	no-
*niman ‘louse’	numan
*n(o,u)man ‘mind, soul’	numan ‘thought, mind, will’

final *n > n

*mundun ‘internal organs’	(?) mundun mo- ‘be pot-bellied’
*niman ‘louse’	numan
*n(o,u)man ‘mind, soul’	numan
cf. also	
*mundu[n]-maŋgV ‘heart’	mund-muŋ

There is one instance each of initial and medial *ŋ > ŋ.

*ŋaŋ[a] ‘baby’	ŋaŋ ‘small male child’

Kalam

Kalam has four nasals: *m*, *n*, *ñ*, *ŋ*. These all occur initially, medially and finally. pTNG *m and *n are regularly reflected as *m* and *n*, respectively, in initial, medial and final positions. In two cases *n may be reflected by a palatal nasal, *ñ*. *ŋ appears as *ŋ* initially and finally.

initial *m > m

*ma- ‘NEG-clitic’	ma-
*ma(n,k,L)[a] ‘ground’	man
*maŋgV ‘compact round object’	magi
*mundun-maŋgV ‘heart’	md-magi (Ti dialect md-magl)
*maŋgat[a] ‘teeth, mouth’	meg
*mapVn ‘liver’	mapn
*(m,mb)elak ‘light, lightning’	melk ‘light’
*mo[k,ŋg]Vm ‘joint’	mogm
*m(o,u)k ‘milk, sap, breast’	muk, mk
*mVn[a]-‘be, live, stay’	md-
*mV ‘taro’	m

*mVkVm ‘jaw, cheek’	mkem ‘cheek’

medial *m > m

*am(a,i) ‘mother’	ami
*kumut, *tumuk ‘thunder’	tumuk
*kumV- ‘die’	kum-
*iman ‘louse’	iman
*n(o,u)man ‘mind, soul’	noman ‘soul’ (poss. borrowed from Chimbu-Waghi)

final *m > m

*kanim ‘cuscus’	kmn ‘game mammal (generic)’ (metath.)
*mo[k,ŋg]Vm ‘joint’	mogm
*mVkVm ‘jaw, cheek’	mkem ‘cheek’
*sVkVm ‘smoke’	skum, sukum

initial *n > n

*n(o,u)man ‘mind, soul’	noman ‘soul’
*na ‘1SG’	-n-, -in ‘1SG subj. agreement’
*nu ‘1PL independent’	-nu-, -un ‘lPL subj agreement’
*nVŋg- ‘know, hear, see’	ng- (Ti dialect), nŋ- ‘see, perceive, etc.’

There are two Kalam forms with initial palatal nasal that may continue eTNG etyma with initial *n- but their cognation is very uncertain.

*nV ‘child’	(?) ñ, ñi
*nok ‘water’	(?) ñg

medial *n > n

*k(a,o)nan ‘shadow’	kawnan ‘spirit of the dead’
*kin(i,u)- ‘sleep’	kn-
*[w]ani ‘who?’	an

final *n > n

*iman ‘louse’	iman
*mapVn ‘liver’	mapn
*takVn[V] ‘moon’	takn
*-Vn ‘1SGsubj agrmt’	-n, -in-
cf. also	
*ma(n,k,L)[a] ‘ground’	man

medial or final *n > ñ

*panV ‘female’	pañ

initial *ŋ > ŋ

*ŋaŋ[a] ‘baby’	-ŋaŋ

medial *ŋ > ŋ

*ŋaŋ[a] ‘baby’	-ŋaŋ ‘baby’

final *ŋ > ŋ

*mbalaŋ ‘flame’	maŋlaŋ, malaŋ
*saŋ ‘story, song’	saŋ ‘women’s song’
*k(o,u)ma(n,ŋ)[V] ‘neck, nape’	koŋam (metath.) (cf. Kobon uŋam, loss of *k regular)

Apalɨ

Apalɨ has three nasals*: m, n, ŋ*. These all occur initially, medially and finally. pTNG *m remains as *m* in Apalɨ in initial and medial positions. *n remains as *n* initially, medially and finally. *ŋ is unattested.

initial *m > m-

*maŋgV ‘compact round object’	maŋgɨ ‘egg’
*mapVn ‘liver’	mapɨn
*maŋgat[a] ‘teeth, mouth’	mɨka
*mVkVm ‘jaw’	mukum
*(m,mb)elak ‘light, lightning’	(Osum and Paynamar mira, Moresada merak)

medial *m > m

*kumV- ‘die’	kɨm-
*k(o,u)ma(n,ŋ)[V] ‘neck, nape’	(sa)kum ‘nape’
*iman ‘louse’	iman

initial *n >n

*na- ‘eat’	n-
pMadang *na ‘2SG free pronoun’	nama (cf. also na- ‘2SG POSS’)
pMadang *nu ‘3SG free pronoun’	numbu (cf. also nu- ‘3SG POSS’)

initial *n > zero

*nVŋg- ‘know, hear, see’	iŋg- (some other S Adelbert languages have niŋg-)

medial *n > n

*kambena ‘arm’	human
*kin(i,u)- ‘sleep’	(?) hɨni- ‘be, stay, exist’

*[w]ani ‘who?’	ani
*(s,nd)umu(n,t)[V] ‘hair’	(?) mɨnɨ

final *n > -n

*mapVn ‘liver’	maßɨn
*iman ‘louse’	iman
*takVn[V] ‘moon’	(Acɨ dial.) takun (Z), (Akɨ dial.) lakun (W)

final *ŋ > ŋ

*sa(ŋg,k)asiŋ ‘sand’	kasɨŋ (Z)

Selepet

Selepet has three nasals *m*, *n*, *ŋ*. Selepet keeps pTNG *m as *m* initially and medially. *n appears as *n* initially and medially. Word-finally, *n is reflected twice as *n*, once as *t*, once as zero. *ŋ is reflected as *ŋ* initially and medially, but becomes *m* finally.

initial *m > m

*masi ‘orphan’	madu
*me(n,t)e ‘head’	mete ‘forehead’
*mV(k,ŋ)V[C] + t(e,i)- ‘vomit’	mohat (cf. Nomu mekat ‘spittle’, Timbe mugat ‘be sick’)
*mVn[a]-‘be, live, stay’	man- ‘live, dwell’
*mo ‘penis’	mɔi

initial *m > b

*(m,mb)elak ‘light, lightning’	belek

medial *m > m

*kumV- ‘die’	(cf. Burum kɔmu-)
*iman ‘louse’	imen
*(s,nd)umu(n,t)[V] ‘hair’	somot
*amu ‘breast’	(n)am ‘breast, milk’
*am(a,i) ‘mother’	(?) (m)ɔmɔ

No reflexes of final *m have been noted.

initial *n > n

*na- ‘eat’	ne-
*nVŋg- ‘know, hear, see’	nɔgɔ- ‘hear, know, listen to s.t.’
*ni ‘1PL free pronoun’	ne(n) ‘1PL’, ne(t) ‘1DL’
*n[e]i ‘bird’	nɔi

medial *n > n

*kan(a,e)ne 'left (side)'	kane
*mVn[a]-'be, live, stay'	man- 'live, dwell'

medial *n > Selepet word-final t

*mbena 'arm'	bɔt

medial *n is lost before i

*kani 'foot'	kɔi

final *n > n

*iman 'louse'	imen
*-Vn '1SGsubj agrmt'	-an '1SG, 1PL'

initial and medial *ŋ > ŋ

*ŋaŋ[a] 'baby'	ŋaŋa

NB. Selepet regularly loses pHuon Peninsula *ŋ, so the retention in *ŋaŋa* may be an exception, as is sometimes the case in nursery words.

final *ŋ > m

*mbalaŋ 'flame'	balam
*mbilaŋ 'tongue'	ni-bilam

In one case a final syllable ending in *ŋ is lost:

*sa(ŋg,k)asiŋ 'sand'	sak

cf. also *kiti(ŋ) 'laugh' > Selepet *giriŋ-nev*, but the final velar nasal in the Selepet form is probably non-etymological.

Kâte

Kâte has three nasals *m*, *n*, *ŋ*. pTNG *m is continued as *m* initially and medially. Likewise, *n remains *n* initially and medially. *ŋ is unattested. In final position all nasals merge as *ŋ*.

initial *m > m

*ma- 'NEG clitic'	mi
*masi 'orphan'	mɔsiŋ
*me(l,n)e 'tongue'	(na)meŋ
*mundun 'internal organs'	munduŋ 'egg'
*mV(k,ŋ)V[C] + t(e,i)- 'vomit'	maŋuzo

initial *m > b

*(m,mb)elak ‘light, lightning’	bɔriʔ ‘glitter, flash of lightning, etc.’

medial *m > m

*amu ‘breast’	ameʔ
*[nd,s]umu[n,t]V ‘hair’	tsiminuŋ ‘stiff coarse hair’
*kumV- ‘die’	hɔmozo
*niman ‘louse’	imeŋ

initial *n > n

*na ‘1SG’	no
*na- ‘eat’	nɔ-
*ni ‘1PL’	ne(n) ‘1PL’, ne(t) ‘2DL’

medial *n > n in one dubious comparison

*kan(a,e)ne ‘left (side)’	(?) kpana

final *n > ŋ

*mundun ‘internal organs’	munduŋ ‘egg’
*iman ‘louse’	imeŋ

Binandere

Binandere (the language, not to be confused with the Binanderean group) has three nasals: *m*, *n*, *ŋ*. pTNG *m remains *m*, initially and medially. *n remains *n* initially and finally. No certain reflexes have been noted of medial *n. *ŋ is not attested. In cases where reflexes of an eTNG etymon have not been found in Binandere, reflexes from other members of the Binanderean group are cited in parentheses.

initial *m > m

*m(i,u)ndu ‘nose’	mendo
*m(o,u)k ‘milk, sap, breast’	mu ‘sap’
*mundun ‘internal organs’	mundu ‘kidney, testicles’
*(ŋg,k)iti-maŋgV ‘eye’	(gisi)-moka ‘eye’
*mV ‘taro’	(Suena ma ‘taro’)
*mV- ‘give’	(Korafe mut- ‘give’)

initial *m > b

*(m,mb)elak ‘light, lightning’	biriga ‘lightning’ (Korafe biria ‘lightning’)

medial *m > m

*am(a,i) ‘mother’	ai (*m lost before i), (Suena mia)
*amu ‘breast’	ami

medial *m > mb

*k(i,u)tuma ‘night, morning’	tumba ‘darkness’ (Suena tumou ‘night’)

initial *n- > n

*na ‘1SG’	na
*na- ‘eat’	na- ‘eat, drink’ (Suena nai- ‘eat’)
*n[e]i ‘bird’	ni
*nVŋg- ‘know, hear, see’	(Korafe niŋg- ‘hear, understand’)
pMadang-Binandere	
*nu[k] ‘3SG free pronoun’	nu

medial *n > n

*ka(m,mb)(a,u)na ‘stone’	ganuma (Korafe ghamana ‘stone’)

final *n is lost

*mundun ‘internal organs’	mundu ‘kidney, testicles’, (Korafe munju ‘egg’)

final is *ŋ is lost in one dubious comparison

*mbalaŋ ‘flame’	(?) beriberi ‘be alight’

Kiwai (Island Kiwai)

Island Kiwai has two nasals *m*, *n*, like all the languages in the Kiwai group. It keeps both *m and *n initially and medially.

initial *m > m

*maŋgat[a] ‘teeth, mouth’	mangota
*m(i,u)ndu ‘nose’	(Gope Kiwai modi, but Island Kiwai wodi)

medial *m > m

*amu ‘breast’	amo
*niman ‘louse’	nimo
*(nd,s)umu(n,t)[V] ‘hair’	? muso (metath?)

initial *n > n

*niman ‘louse’	nimo
*ni ‘1PL’	ni(mo)

medial *n > n

*takVn[V] ‘moon’	sagana
*mbena ‘arm’	(Kerewo Kiwai bena ‘shoulder’)

Telefol

Telefol has three nasals. *m* and *n* occur initially medially and finally. *ŋ* occurs only medially and finally. pTNG *m are retained initially, medially and finally. *n is continued initially and medially. *ŋ appears as ŋ, finally, in its sole attestation.

initial *m > m

*m(o,u)k 'milk, sap, breast'	müük, mɔk 'spittle'
*maŋgat[a] 'teeth, mouth'	(Faiwol makat-kalim 'whiskers')
*maŋgV 'compact round object'	magap 'round object, fruit, seed, etc''
*m(i,u)ndu 'nose'	mutu 'nose'

medial m > m

*k(o,u)ma(n,ŋ)[V] 'neck, nape'	kum 'left side of neck'
*kumut, *tumuk 'thunder'	tumuun 'thunder'
*niman 'louse'	tim 'louse'

final m > m

*kal(a,i)m 'moon'	kaliim 'moon'
*k(i,u)tuma 'night, morning'	kutim 'morning'

initial *n > n

*na '1SG'	na-
*ni, *nu '1PL'	nu

medial *n > n

*mbena 'arm'	ban 'forearm'
*[w]ani 'who?'	wan(tap), waan(ta) 'who?'
*pVnum 'wind'	inim

medial *n > n

*kinV 'shoulder'	tiŋ (Faiwal kiiŋ)

final *ŋ > ŋ

*mbilaŋ 'tongue'	fɔŋ (cf. Faiwol falaŋ, Tifal filaŋ)

Kaeti

Kaeti has two nasals *m*, *n*. It keeps pTNG *m initially. Medially, one reflex shows *m*, another *n*, but cognate forms in sister Awyu-Dumut languages show *m* in both cases. Final *m is lost in the single attested instance. *n is retained as *n* initially but not attested medially. *ŋ is not attested.

initial *m > m

*maŋgat[a] 'teeth, mouth'	magot
*mVkVm 'cheek'	(a)moka (cf. Axu moxo pe)
*maŋgV 'compact round object'	(Axu mügo 'egg')

medial *m > m

*amu 'breast'	am
*k(o,u)ma(n,ŋ)[V] 'neck, nape'	koman

medial *m > n

*kumV- 'die'	kün (cf. Sawuy xom-, Wambon N. & Wambon S. kim-)

final *m is lost

*mVkVm 'cheek, jaw'	(a)moka 'cheek' (cf. Axu moxo pe 'cheek')

initial *n > n

*na '1SG'	nø(p)
*ni, *nu '1PL free pron.'	no-güp
*na- 'eat'	(Wambon en-)

final *n or *ŋ > n

k(o,u)ma(n,ŋ)[V] 'neck, nape'	koman

Asmat (Flamingo Bay Asmat)

The Flamingo Bay Asmat sound system is analysed by Voorhoeve (1965) as having 11 consonant phonemes: /p t c k f s m n w r j/. /m/ has allophones [b] initially, [mb] before a nasal and [m] elsewhere. /n/, likewise, has allophones [d] initially, [nd] before a nasal and [n] elsewhere. There are few reflexes of pTNG nasals in Asmat but in these *m and *n generally remain as *m* and *n*.

initial *m > m

*maŋgat[a] 'teeth, mouth'	me
*(m,mb)elak 'light, lightning'	mer

medial *m > m

*niman 'louse'	(Kamoro namo)

initial *n > n

*na- 'eat'	na-
*ni, *nu 'IPL'	na '1PLinc', na(r) '1PLexc'

medial *n > n
*mun(a,i,u)ka ‘egg’ manaka

final *n > n
*niman ‘louse’ (cf. Kamoro namo)

final *n > zero
*kasin ‘mosquito’ isi

4.3.2.3 Obstruents

The TNG obstruents present a far more difficult challenge than the nasals. A first comparison of the sample languages realises scores if not hundreds of partially different correspondence sets. One can make a good deal of sense of these by assuming that:
(1) pTNG had two contrasting sets of obstruents: a prenasalised and voiced set and an oral (and probably voiceless in at least initial and final positions) set, and that in each case there was a contrast between bilabial, apical and velar points of obstruction: *mb vs. *p, *nd vs. *t, and *ŋg vs. *k. There was also an alveolar oral fricative, *s, and possibly a prenasalised alveolar or palatal affricate, *nj.
(2) These obstruents showed a good deal of phonetically conditioned allophonic variation that favoured phonetic change. For example, given the patterns of allophonic variation seen in contemporary TNG languages it is quite likely that the oral obstruents *p and *k had voiced fricative allophones between vowels and that *t may have been realised as a tap [r] between vowels and as an affricate [[tj] or [ts] before front vowels.
(3) In any case, phonetically conditioned sound changes along these lines occurred many times independently in daughter languages.
(4) Certain kinds of irregular (lexically sporadic) sound changes occurred from time to time, including replacement of prenasalised stops by homorganic nasals, and vice versa, e.g. mb > m, m > mb, nd > n, n > nd.

The most unsatisfactory part of the obstruent system as first reconstructed concerned *nd. This symbol subsumed many partially distinct correspondence sets. The next step was to group the various sets into two main classes, represented by *nd and *nj, where *nj subsumes sets where a high proportion of languages show affricative or fricative reflexes, with the rest assigned to *nd. This distinction still leaves much unexplained diversity among correspondence sets, especially those assigned to *nj. It is unfortunate that *nj reflexes are poorly represented in the sample of languages used here to demonstrate reflexes of pTNG consonants.

Table 6 shows the frequencies of obstruents in a corpus of 190 eTNG etyma. Where one segment of an etymon is reconstructed with alternative obstruents, e.g. *nj or *s, in *ke(nj,s)a ‘blood’, both alternatives are counted.

Table 6 Frequencies of obstruents in eTNG lexical reconstructions

	*mb	*nd	*nj	*ŋg	*p	*t	*s	*k
initial	7	3	1	4	4	7	5	30
medial	10	6	5	7	4	22	4	4
final	–	–	–	–	1	2	–	8

Certain obstruents appear far more frequently in reconstructions than others and certain obstruents appear as favouring certain positions in the word. *k is by far the most frequent obstruent word initially and *t the most frequent medially. Prenasalised obstruents are not attested in word final position. They are somewhat more frequent medially than initially.

Table 7 summarises the most common reflexes of the obstruents in the sample languages. A dash indicates that no reflex is attested. This table represents a considerable simplification of the full range of reflexes. In the language-by-language exemplification that follows, it will be seen that many of the correspondences are attested by just one or two reflexes and that in some cases there are multiple correspondences with no obvious conditioning factors.

Table 7. Reflexes of pTNG obstruents in a sample of 10 languages

pTNG	**mb-**	**-mb-**	**nd-**	**-nd-**	**ŋg-**	**-ŋg-**
M. Wahgi	–	mb	–	nd	–	ŋg, ŋ
Kalam	mb	mb	–	nd	–	ŋg
Apalɨ (Akɨ)	–	mb	–	nj	–	ŋg
Selepet	b	b,p	s,t	–	g	g
Kâte	b	mb,p	s,t	s	g	k
Binandere	–	p, mb	–	nd,z	g	k
Kiwai	b,p	p	–	d,t	–	g
Telefol	f	b	–	n	–	k
Kaeti	b	–	d	d	–	g
Asmat	mb	p	–	–	–	k

pTNG	**p-**	**-p-**	**-p**	**t-**	**-t-**	**-t**	**k-**	**-k-**	**-k**
M. Wahgi	p	p	–	t	–	–	k	k	k
Kalam	p	p	–	t	–	t, Ø	k	k	–
Apalɨ	–	β	Ø	l,t	t	Ø	h,k	h,k	–
Selepet	–	–	p	t	r	t(?)	k	k	k
Kâte	f	f	t	t	–	-	k,h	–	ʔ
Binandere	–	–	–	t,j[(1)]	r,s/_i	–	k	k	-
Kiwai	–	–	–	s,t	r,t	t	g, Ø	g	–
Telefol	f	–	–	t	t	–	k	k	k
Kaeti	b	–	–	–	t	t	k	–	k
Asmat	f(?)	–	t,r	s[(2)]	r,s,t	–	Ø	k	–

pTNG	**s-**	**-s-**	**nj-**	**-nj-**
M. Wahgi	–	–	–	–
Kalam	s	s	–	–
Apalɨ	s	s	–	–
Selepet	t, s	s,d	–	nd
Kâte	s	–	–	–
Binandere	s	s	s(?)	z
Kiwai	s,t(?)	–	–	r
Telefol	s(?)	–	–	–
Kaeti	–	–	–	–
Asmat	t,s	s	–	s

NOTES: (1) *t > j before i in Binandere. (2) *t > s before i in Asmat.

The overall correspondence sets for the obstruents offer little evidence for high order subgroupings.

Middle Wahgi

Middle Wahgi distinguishes four prenalised obstruents /mb, nd, nj, ŋg/ and four oral obstruents /p, t, s, k/. /t/ is realised as flapped [ř] medially. /p/ and /k/ are voiced medially, /s/ is [ts] initially, [s] medially and finally it may be any of [c,s,z].

Middle Wahgi appears to keep separate reflexes of seven pTNG obstruents (there are no reflexes of *nj) in at least some contexts. However, in most cases only a single reflex has been found.

medial *mb > mb

*ambi ‘man’	(?) amb ‘woman’, ambi- ‘wife’
*imbi ‘name’	embe(m)

initial *p > p

*pu- ‘go’	pu-

medial *p > p

*apa ‘father’	apa- ‘maternal uncle’

medial *nd > nd

*mund-mangV ‘heart’	mund-mung

initial *t > t

*tVk- ‘cut, cut off’	tuk- ‘chop’

medial *ŋg > -ŋg
*maŋgV ‘compact round object’ mungum ‘kidney’

medial *ŋg > M. Wahgi word-final ŋ
*maŋgV ‘compact round object’ muŋ ‘fruit, nut, lump’
*mundu[n]-maŋgV ‘heart’ mund-muŋ

(?) initial *k > k
*kakV- ‘carry on shoulder’ (?) kau- ‘carry on head or shoulder’

medial or final *k > k
*tVk- ‘cut, cut off’ tuk- ‘chop’
*muk ‘blue’ muk
*mV(k,ŋ)V[C] + t(e,i)- ‘vomit’ mek (si-), mek ‘vomitus’

Kalam
Kalam has four prenalised obstruents /b, d, j, g/ pronounced [mb, nd, ñǰ, ŋg] word initially and medially and [mp, nt, ñč, ŋk] word finally. It has five oral obstruents /p, t, s, c, k/, most with two or more allophones. Phonemic consonant clusters within words are separated by an epenthetic vowel, usually a short high central vowel, e.g. *bsg* ‘sit’ [mbɨsɨŋk] and when roots consisting of a single consonant are pronounced alone there is a release vowel, e.g. /b/ ‘man’ [mbə]. Kalam has no reflexes of *nj but retains distinct reflexes of the other seven pTNG obstruents.

initial *mb > b
*mbapa ‘father’ bapi

initial *mb > m in at least one case
*mbalaŋ ‘flame’ malaŋ, maŋlaŋ

medial *mb> b
*ambi ‘man’ b [mbə]
*imbi ‘name’ yb [yimp]
*ka(m,mb)(a,u)na ‘stone’ kab [kamp]
*sambV ‘cloud’ seb [semp]
*simb(i,u) ‘guts’ sb [sɨmp]

***p > p initially and medially** (Kalam /p/ is realised as [ɸ] initially, [β] medially)
*panV ‘female’ pañ
*apus[i] ‘grandparent’ aps
*mapVn ‘liver’ mapn

medial *nd > d

*kindil 'root'	kdl
*mundu[n]-maŋgV 'heart'	mudmagi

initial *t > t (Kalam /t/ is [t] initially, [r] elsewhere)

*takVn[V] 'moon'	takn
*tVk- 'cut, cut off'	tk- 'sever, cut off'
*tu 'axe'	tu
*kumut, *tumuk 'thunder'	tumuk

final * t > t

*-i(t,l) '2DL verbal suffix'	-it

medial or final *t > zero in case of syllable loss

*maŋgat[V] 'teeth, mouth'	meg

***s > s initially and medially**

*sambV 'cloud'	seb
*simb(i,u) 'guts'	sb
(?)*su- 'bite'	su-
*apus[i] 'grandparent'	aps 'grandmother'

medial *ŋg > g

*maŋgat[a] 'teeth, mouth'	meg
*maŋgV 'compact round object'	magi

In one verb root medial *ŋg has varying reflexes in different dialects of Kalam. *nVŋg- 'know, hear, see' ng- or nŋ- in Ti dialect, but only nŋ- in Etp dialect.

initial *k > k in all positions (Kalam /k/ is realised as [ɣ] medially, [k] elsewhere)

*ka(m,mb)(a,u)na 'stone'	kab
*k(a,o)nan 'shadow'	kawnan
*kin(i,u)- 'sleep'	kn-
*kumV- 'die'	kum-
*k(o,u)ma(n,ŋ)[V] 'neck, nape'	koŋam (metath.)
*kakV- 'carry on shoulder'	kak-
*m(o,u)k 'milk, sap, breast'	muk (Ti dial. mok)
*muk 'brain'	muk
*takVn[V] 'moon'	takn
*tVk- 'cut, cut off'	tk- 'sever'

Apalɨ

There are two main dialects, Acɨ and Akɨ, which differ significantly. Akɨ has eight oral obstruents: f, p, v, t, c, s, k, h, where /h/ is a voiced velar fricative [ɣ] with a voiceless allophone [x] (conditions not specified). However, /p,t,k/ are rare in Akɨ. Where Acɨ has /p,t,k/, Akɨ often has /v,l,h/ respectively. Both dialects have four prenasalised obstruents: b, d, j, g, which I will write here as *mb, nd, nj, ŋg*. Prenasalised obstruents do not occur word initially or finally, and so are open to reanalysis as phonemic clusters.

Wade (n.d.) describes the Akɨ dialect. Zgraggen (1980d) appears to have recorded Acɨ. In items cited below these two sources are marked (W) and (Z). For Zgraggen's use of barred u for the high central vowel, I have substituted ɨ, in accordance with Wade's phonological description.

*nj is not attested in Apalɨ but this language retains distinct reflexes of the other seven pTNG obstruents.

initial *mb > medial p

*mb(i,u)t(i,u)C 'fingernail'	tɨpi (metath.) (Z)

medial *mb > mb

*imbi 'name'	imbi (W)
*[ka]tumba(C) 'short'	tɨmbɨ (W)
*si(mb,p)at[V] 'saliva'	sɨmbu 'spit' (W), simbiŋ (Z)
*simbil[VC] 'navel'	(Akɨ) simbilɨm, (Acɨ) cimbilɨm 'placenta, navel, umbilical cord' (W)
*si(m,mb)(i,u) + modifier 'buttocks'	susum 'lower buttocks' (W)

medial *mb > m

*kambena 'arm, forearm'	human (W)

medial *p > β

*mapVn 'liver'	maβɨn (W)
*apa 'father'	iaβaŋ (W)
*apus[i]'grandparent'	aβe 'grandmother' (W)

(?) medial *p > f

*apa(pa)ta 'butterfly'	(?) afafaŋ (Z)

medial *nd > nj

*ka(nd,t)(e,i)kV 'ear'	hinji (W)
*kindil 'root'	hɨnjilɨ (W)

initial *t > t

*[ka]tumba(C) ‘short’	tɨmbɨ (W, Z)
*takVn[V] ‘moon’	(Acɨ dial.) takun (Z)

initial *t > l (in Akɨ dialect)

*takVn[V] ‘moon’	lakun (W)
*kumut, *tumuk ‘thunder’	lɨmbɨ(lami) ‘to thunder’ (W)

medial *t > t

*mb(i,u)t(i,u)C ‘fingernail’	tɨpi (metath.) (Z)
*kit(i,u) ‘leg’	gɨtɨ (Z)
*kutV(mb,p)(a,u)[C] ‘long’	(Akɨ) hutaŋ (W), (Acɨ) kutes (Z)

medial or final *t is lost

*si(mb,p)at[V] ‘saliva’	sɨmbu ‘spit’ (W), simbiŋ (Z)
*maŋgat[a] ‘teeth, mouth’	mɨka (W)

initial *s > s

*si(mb,p)at[V] ‘saliva’	sɨmbu ‘spit’ (W), simbiŋ (Z)
*simb(i,u) ‘guts’	su ‘faeces’ (W)
*simbil[VC] ‘navel’	(Akɨ) simbilɨm, (Acɨ) cimbilɨm (both W)
*si(m,mb)(i,u) + modifier ‘buttocks’	susum (W)

medial *s > s

*sa(ŋg,k)asiŋ ‘sand’	kasɨŋ (Z)

medial *ŋg > ŋg

*maŋgV ‘compact round object’	maŋgɨ ‘egg’ (W)
*nVŋg- ‘know, hear, see’ i	ŋg- ‘see’ (W)

medial *ŋg > k

*maŋgat[a] ‘teeth, mouth’	mɨka (W)

initial *k > k

*kumV- ‘die’	kɨm- ‘die’ (Z), kɨmukɨmu ‘dry, dead’ (W)
*k(o,u)ma(n,ŋ)[V] ‘neck, nape’	(sa)kum ‘nape’ (W)

initial *k > h

*kambena ‘arm’	human (W)
*kindil ‘root’	hɨnjɨlɨ (W) (cf. gɨndrɨ ‘root’ (Z))
*ka(nd,t)(e,i)kV ‘ear’	hɨnji (W)
*kin(i,u)- ‘sleep, lie down’	hɨni- ‘be, stay, exist’ (W)
*kutV(mb,p)(a,u)[C] ‘long’	hutaŋ (W)

*kumV- ‘die’	hɨmi- (W)

initial *k > g

*kit(i,u) ‘leg’	gɨtɨ (Z)

medial *k > h in Acɨ, k in Akɨ

*mVkVm ‘cheek’	(Acɨ) mukum (W), Akɨ mɨhum (W)

medial *k > k

*takVn[V] ‘moon’	lakun (W)

medial *k is lost

*ka(nd,t)(e,i)kV ‘ear’	hɨnji (W)

final *k is lost in one dubious comparison

*tumuk ‘to thunder’	(?) lɨmbɨ(lami) (W)

Selepet

Selepet has eight obstruents. There are three voiceless stops: p, t, k, three voiced stops: b, d, g, and two fricatives: s and h. Selepet possibly merges *nj and *s. It appears to preserve distinct reflexes of all other obstruents in at least some positions. However, there are multiple reflexes of some consonants, e.g. of *k and *s, with no known conditions.

initial *mb > b

*mbalaŋ ‘flame’	balam
*mbilaŋ ‘tongue’	[ni]bilim
*mbena ‘arm’	(?) bot

initial *b > p

*mbulikV ‘turn (oneself)’	purik (sɔ-)

Selepet *purik* is a verbal adjunct paired with the inflecting verb *sɔ-*.

medial *mb > b

*imbi ‘name’	ibi
*sambV ‘cloud’	hibim ‘sky’ (Kâte sabɔŋ, Komba sumbem ‘sky’)
*ambi ‘man’	(?) ibi ‘woman’ (Burum ambi ‘woman’)

medial *mb > Selepet word final p

*simb(i,u) ‘guts’	tep- ‘stomach, intestines’
*si(mb,p)at[V] ‘saliva’	(Sialum sawat, Migabac sofoʔ, Nomu sowot)
*kutV(mb,p)(a,u)[C] ‘long’	kɔlip

medial *mb is lost

*kambu 'ashes'	kɔu

final *p > p

*kend(o,u)p 'fire'	kɔlɔp

initial nd > s

*nde- 'speak'	sɔ-

medial *nd > l

*kend(o,u)p 'fire'	kɔlɔp

medial *nd > nd in one dubious comparison

*ka(nd,t)(e,i)kV 'ear'	(?) ɔndɔp (prob. a loanword. Suter (p.c.) reconstructs pHuon Peninsula *kazap 'ear')

initial *t > t

*tVk- 'cut, cut off'	tok [yap] 'snap, break'

medial *t > l

*kutV(mb,p)(a,u)[C] 'long'	kɔlip

medial *t > r

*kVtak 'new'	irak

final *t > t

*(s,nd)umu(n,t)[V] 'hair'	somot

initial *s > s

*sa(ŋg,k)asin 'sand'	sak
*(s,nd)umu(n,t)[V] 'hair'	somot
*si(mb,p)at[V] 'saliva'	(Sialum sawat, Migabac sofoʔ, Nomu sowot)

initial *s > t in one dubious comparison

*simb(i,u) 'guts'	(?) tep- 'stomach, intestines'

initial *s > h

sambV 'cloud'	hibim 'sky' (cf. Kâte sambɔŋ, Komba sumbem 'sky')

medial *s > d

*masi 'orphan'	madu

initial *nj > s

*nj(a,e,i)- 'burn'	si-

initial and medial *ŋg > g

*ŋga '2SG'	ga
*nVŋg- 'know, hear, see'	nɔgɔ- 'hear, know, listen to s.t.'

initial *k > k

*kend(o,u)p 'fire'	kɔlɔp
*kakV- 'carry on shoulder'	kaku-
*kutV(mb,p)(a,u)[C] 'long'	kalip
*kambu 'ashes'	kɔu

medial *k > k

*kakV- 'carry on shoulder'	kaku-
*mbulikV 'turn (oneself)'	purik
*sa(ŋg,k)asin 'sand'	sak

medial *k > h

*mV(k,ŋ)V[C] + t(e,i)- 'vomit'	mohat (Nomu mekat 'spittle', Timbe mugat 'be sick')

final *k > k

*kVtak 'new'	irak
(m,mb)elak 'light, lightning'	belek 'lightning'

medial or final *k > k in one dubious comparison

*ok[V] 'water'	(?) (gel-)ok 'rain'

Kâte

Kâte has 13 obstruents. There are five voiceless stops: p, t, k, kp, ʔ, four voiced stops: b, d, g, gb, and four fricatives/affricates: s, ts, dz and h. kp and gb are coarticulated bilabial-velar stops. Kâte keeps distinct reflexes of *mb and *p, and *nd and *t. Evidence about the outcome of other pairs of pTNG prenasalised and oral obstruents is less clear, but it suggests that the pTNG contrasts were maintained between most pairs.

initial *mb > b

*mbalaŋ 'flame'	bɔruŋ

*mb(i,u)t(i,u)C ‘fingernail’	bʉtoŋ
*mbeŋga(-masi) ‘orphan’ ‘widow and child’	bekɔ

medial *mb > mb

*sambV ‘cloud’	sambɔŋ ‘sky’

initial *mb > m

*mbena ‘arm’	me

initial *p > f in one dubious comparison

*(mb,p)ututu- ‘to fly’	(?) fururuʔ

medial *-mb/p > f

*si(mb,p)at[V] ‘saliva’	tofeʔ

initial *t > t

*tVk- ‘cut, cut off’	tɔʔ(ne)

initial *nd or *t > l

*(nd,t)a- ‘take’	lo-

medial *nd > nd

*mundun ‘internal organs’	mʉnduŋ ‘inner yolk of egg’

medial *t > r in one dubious comparison

*(ŋg,k)atata ‘dry’	(?) kereŋke-

initial and medial *s > s

*sambV ‘cloud’	sambɔŋ ‘sky’
*masi ‘widow’	masiŋ

initial *s > t in one dubious comparison

*si(mb,p)at[V] ‘saliva’	(?) tofeʔ

initial *ŋg- > g

*ŋga ‘2SG’	go

medial *ŋg > k

*mbeŋga(-masi) ‘orphan’, ‘widow and child’	bekɔ

initial *k > h

*kumV- ‘die’	hɔmo-

initial *k > k in two dubious comparisons

*ka(m,mb)(a,u)na ‘stone’	(?) kpana
*kV(mb,p)(i,u)t(i,u) ‘head’	(?) kpit(seʔ)

final *k > ʔ

*(m,mb)elak ‘light, lightning’	bɔriʔ ‘glitter, flash of lightning, etc.’

Binandere

For Binandere there are some indications that a contrast is kept between most pairs of pTNG obstruents. However, the data are too skimpy and inconsistent to be conclusive about all such contrasts.

initial *mb > b

*mbalaŋ ‘flame’	beri-beri ‘be alight’
*mbulikV ‘turn (oneself)’	(Guhu-Samane burisi eetaqu ‘turn over, turn s.th. around’)
*mbeŋga-masi ‘orphan’, ‘widow and child’	(Suena boga masa ‘destitute widow and child’

initial *p > p

*pu + verb ‘to blow’	Binandere put- ‘blow’

medial *mb > mb

*ambi ‘man’	embo (Guhu-Samane abi ‘man’)

medial *mb/p > p

*kV(mb,p)(i,u)t(i,u) ‘head’	kopuru
*[ka]tumba(C) ‘short’	tupo
*kambu(s,t)(a,u) ‘smoke’	(?) imbosi

medial *p > f

*apa ‘father’	afa (Korafe afa)

initial *nd > d

*ndaŋgi/ndiŋga ‘tie’	(Suena di ‘tie’)

medial *nd > nd

*m(i,u)ndu ‘nose’	mendo
*mundun ‘internal organs’	mundu ‘kidney, testicles’ (Korafe munju ‘egg, etc.’)

medial *nd > r

*ka(nd,t)(e,i)kV ‘ear’	(Yega kari ‘ear’)

medial *nd > z

*inda ‘tree’	izi (cf. Notu ri)

initial *t > t

*[ka]tumba(C) ‘short’	tupo

medial *t > initial t

*k(i,u)tuma ‘night, morning’	Binandere tumba ‘darkness’ (Suena tumou ‘night’)

medial *t > r

*kV(mb,p)(i,u)t(i,u) ‘head’	kopuru

medial *t > s / _ i

*(ŋg,k)iti-maŋgV ‘eye’	gisi moka
*at(i,u) ‘netbag’	asi (Suena ati ‘netbag’)

initial *s > s

*si[si] ‘urine’	pBinandere *susu (Korafe soso)

initial *t > j before i

*titi ‘tooth’	ji

medial *s > s

*asi ‘string, rope’	asi ‘vine, string, rope’)
*kasipa ‘to spit’	kosiwa ‘spittle’
*mbeŋga-masi ‘orphan’, ‘widow and child’	(Suena boga masa ‘destitute’)

(?) medial *nj > r in one dubious comparison

*kanjipa ‘sun’	(?) kariga ‘moon’ (-g- unexp.)

initial *ŋg > g

*(ŋg,k)iti-maŋgV ‘eye’	gisi-(moka)

medial *ŋg > k

*(ŋg,k)iti-maŋgV ‘eye’	(gisi)-moka (Korafe móko ‘core, centre’)

medial *ŋg > ŋg

*nVŋg- ‘know, hear, see’	(pBinandere *niŋg- ‘hear’, Korafe niŋg- ‘hear, understand’)
*mbeŋga-masi ‘orphan’,	(Suena boga-masa ‘destitute’)

'widow and child'

initial *k > k

*kV(mb,p)(i,u)t(i,u) 'head'	kopuru
*ka(nd,t)(e,i)kV 'ear'	(Yega kari)
*kasipa 'to spit'	kosiwa 'spittle', kosiwa ari 'to spit'
*ka(m,mb)(a,u)na 'stone'	ganuma (metath.) (Korafe ɣamana)

initial *k > g

*ka(m,mb)(a,u)na 'stone'	ganuma (metath.) (Korafe ɣamana 'stone')

initial *k is lost in three dubious comparisons

*k(o,u)ndVC 'bone'	(?) undoru 'bones'
*kumV- 'die'	(?) abu-bugari 'dead people', (pBin *ambu- 'wither, be sick, dying')
*kambu(s,t)(a,u) 'smoke'	(?) imbosi

medial *k > k

*ka(nd,t)(e,i)kV 'ear'	(Yega kari)
*la(ŋg,k)a 'ashes'	(aßa)-raka 'fire' (Korafe aßa-raka 'burning stick')

medial *k > ŋg in one dubious comparison

*sikal/*sakil 'hand, claw'	(?) siŋgu 'finger' (Guhu-Samane sika 'little finger')

final *k > g

*(m,mb)elak 'light, lightning'	biriga 'lightning' (Korafe biria 'lightning')

Kiwai (Island Kiwai)

Kiwai has four voiceless stops /p, t, k, ʔ/, three voiced stops /b, d, g/ and a single fricative /s/. Most languages in the Kiwaian group have the same inventory of consonants. Kiwai appears to keep distinct reflexes of pTNG *mb, *nd, *nj, *t, *k and *s, in some cases. On the slight evidence available, Kiwai appears to have merged *mb and *p, and *ŋg and *k.

initial *mb > b

*mbena 'arm'	(Kerewo bena 'shoulder')

initial *mb > p

*mb(i,u)t(i,u)C 'fingernail'	pitu

medial *mb > p

*kV(mb,p)(i,u)t(i,u) 'head'	epuru, (Wabuda kepuru)
*tukumba[C] 'short'	(?) kopu
*a(mb,m)u 'tail'	(?) wapo

initial *p > medial b

*pi(n,nd)a 'sister'	abida

medial *nd > d

*m(i,u)ndu 'nose'	wodi (Gope (N.E. Kiwai) modi)
*pi(n,nd)a 'sister'	abida

medial *nd > t

*k(a,o]ndok[V] 'foot'	Gope (N.E. Kiwai) oto, Morigi kota

medial *nd > r

*ka(nd,t)(e,i)kV 'ear'	gare

medial *nj > r

*inja 'tree, wood, fire'	(S. Kiwai era)

medial *t > t

*mb(i,u)t(i,u)C 'fingernail'	pitu

medial *t > r

*kV(mb,p)(i,u)t(i,u) 'head'	epuru (Wabuda kepuru)

final *t > t

*maŋgat[a] 'teeth, mouth'	magata

initial *s > s in one dubious comparison

*(nd,s)umu(n,t)[V] 'hair'	?muso (metath)

(?) initial *s > t

*sumbu 'white ashes'	tuwo

medial *s is unattested

initial *k > k

*k(a,o]ndok[V] 'foot'	(Morigi kota)
*kV(mb,p)(i,u)t(i,u) 'head'	(Wabuda kepuru)
*kuk(a,u)m(o,u) 'cold'	(Bamu kukamu, Sisiame kukamo)

initial *k > zero

*kV(mb,p)(i,u)t(i,u) 'head'	epuru
*k(a,o]ndok[V] 'foot'	(Gope (N.E. Kiwai) oto)

initial *k > g

*ka(nd,t)(e,i)kV 'ear'	gare (Bamu Kiwai gare)

medial *k > g

*takVn[V] 'moon'	sagana

medial *ŋg > g

*maŋgat[a] 'teeth, mouth'	magata

Telefol

Telefol has five obstruents: /b, d, t, k, kw/. /b/ is a voiced fricative [β] medially, a voiceless stop [p] finally. There is no contrast between [p] and [b] or between [k] and [g]. The voiced stop /d/ does not occur finally. Telefol appears to merge pTNG *nd and *t as *t*, and *ŋg and *k as *k*, but possibly keeps separate reflexes of *mb vs *p, and of *s vs *nd and *t.

medial *mb > b

*mbena 'arm'	ban 'forearm'
*amba 'sibling'	baab
*(kambu)-sumbu 'ashes'	(ku)-tab

initial *mb > f

*mbilaŋ 'tongue'	foŋ (Tifal filaŋ)

(?) initial *p > f in one dubious comparison

*(mb,p)ututu- 'to fly'	(?) fúlúluú (+ V.)

(?) initial *p is lost in one dubious comparison

*pVnum 'wind'	(?) inim

medial *nd > t

*m(i,u)ndu 'nose'	mutuum

***t > t initially and medially**

*kumut, *tumuk 'thunder'	tumuun
*k(i,u)tuma 'night, morning'	kutim

medial *t > t

*ŋgatu(k,n) 'knee'	katuun

medial *nd > t

*k(a,e)(nd,t)ak 'neck'	ditak (Faiwal getak)

***s > s initially**

*saŋ 'story, song'	saŋ 'myth, story'

initial *s > medial t

*sumbu 'ashes'	(ku-)tab (ku- probably reflects the first morpheme in an old compound; see 5.5)

***ŋg > k medially**

*maŋgV 'compact round object'	(úún) makáb 'egg'
*maŋgat[a] 'teeth, mouth'	(Faiwal makat-kalim 'whiskers (lit. chin-hair)')

***k > k initially**

*kal(a,i)m 'moon'	kaliim
*k(o,u)ma(n,ŋ)[V] 'neck, nape'	kum 'left side of neck'
*k(o,u)ndVC 'bone'	kun
*kutV(mb,p)(a,u)[C] 'long'	(Kati M. kudub)

***k > t initially**

*kinV 'shoulder'	tiŋ-

***k > k finally**

*m(o,u)k 'milk, sap, breast'	múúk
*ok[V] 'water'	óók

***ŋg/k > k initially**

*(ŋg,k)a(nd,t)apu 'skin, bark'	káál

Kaeti

Kaeti has three oral obstruents /p, t, k/ and three matching prenasalised obstruents /mb, nd, ŋg/. There are no fricatives or affricates. Kaeti appears to have kept apart pTNG *nd vs *t and *ŋg vs *k. The situation regarding *mb and *p is less clear.

initial *mb > b

*mb(i,u)t(i,u)C 'fingernail'	betit

medial *mb > p

*imbi 'name'	üp

initial *p is unattested

medial *p > p

*apa[pa]ta 'butterfly'	apap

medial *nd > d

*k(a,o)ndok[V] 'foot, leg'	kodok
*andu- 'to cook'	odu

medial *t > t

*mb(i,u)t(i,u)C 'fingernail'	betit

medial *t > r in two dubious comparisons

*(ŋg,k)iti-maŋgV 'eye'	(?) kerop
*(mb,p)ututu- 'to fly'	(?) bere(na)

final *t > t

*kumut, *tumuk 'thunder'	komöt
*maŋgat[a] 'teeth, mouth'	magot

initial *ŋg > g

*ŋga '2SG'	gu

medial *ŋg > g

*maŋgat[a] 'teeth, mouth'	magot
*maŋgV 'compact round object'	(Axu mügo 'egg')

initial *k > k

*ka(nd,t)(e,i)kV 'ear'	kere(top)
*k(a,o)ndok[V] 'foot, leg'	kodok
*ka(nd,t)apu 'skin'	kotae
*kumbutu 'wind'	kiow
*kin(i,u)- 'sleep'	kinum
*kumV- 'die'	kün
*k(o,u)ma(n,ŋ)[V] 'neck'	koman
*kuya 'cassowary'	(Sawuy kuye)

(?) initial *k > x in one dubious comparison

*kV(mb,p)(i,u)t(i,u) 'head'	(?) xebia(an)

medial *k > k

*mVkVm 'cheek'	(a)moka (cf. Axu moxo pe)

medial *k > g in one dubious comparison

*kutV(mb,p)(a,u)[C] 'long'	(?) guru(op)

medial or final *k > final k

*ok[V] 'water'	ok
*k(a,o)ndok[V] 'foot'	kodok

*s is not attested in any clear cases.

Asmat

Asmat /m/ has allophones [b] initially, [mb] before a nasal and [m] elsewhere. /n/, likewise, has allophones [d] initially, [nd] before a nasal and [n] elsewhere. The evidence is too slight to determine whether Asmat maintained distinct reflexes of *mb vs *p, and *nd vs *t. *k is lost initially in three of four attestations. Medially, *t appears to have merged with *s and *nj.

initial *mb > [b]

*mbena 'arm'	man [ban]

initial *mb > f

*mb(i,u)t(i,u)C 'fingernail'	fit

medial *mb> p

*imbi 'name'	yipi

medial *mb/p > p

*si(mb,p)at[V] 'saliva'	(me)sep
*(mb,p)ututu- 'to fly'	(?) pi-

medial *mb/p > [w] /u_u

*kV(mb,p)(i,u)t(i,u) 'head'	kuwus

medial *nd > s

*inda 'fire'	(Central Coast Asmat isi)

initial *t > t

*tututu[ku] 'straight'	toror

medial *t > r

*k(i,u)tuma 'night, morning'	iram

*tututu[ku] ‘straight’	toror

initial *t > j /_i

*ti, *titi ‘tooth’	ji

medial *t > s

*ata ‘excrement’	asa
*(ŋg,k)atata ‘dry’	soso
*kV(mb,p)(i,u)t(i,u) ‘head’	kuwus

medial *s > s

*kasin ‘mosquito’	(Citak Asmat isi)

medial *nd > s

*inda ‘fire’	(Central Coast Asmat isi)

medial *nj > s

*ke(nj,s)a ‘blood’	es

medial *ŋg > k

*maŋgV ‘compact round object’	moko-per ‘navel’
*mun(a,i,u)ka ‘egg’	manaka

initial *k is lost

*ke(nj,s)a ‘blood’	es
*kasin ‘mosquito’	(Central Asmat isi)
*k(i,u)tuma ‘night, morning’	yiram

initial *k > k

*kV(mb,p)(i,u)t(i,u) ‘head’	kuwus

final *k lost as part of final syllable loss

*(m,mb)elak ‘light, lightning’	(Flamingo Bay Asmat mer ‘lightning’)

4.3.2.4 Laterals

There is a small but fairly convincing set of correspondences supporting the reconstruction of *l, probably an alveolar lateral, possibly retroflexed and/or flapped (see 4.1). There is some evidence, much more problematic, for a second lateral, which will be written *L here.

pTNG *l is attested initially in two reconstructions: *la[ŋ,k]a ‘ashes’, *li- ‘to do’

*l is attested medially in at least seven eTNG etyma: *kal(a,i)m ‘moon’, *kamali ‘sun’ (this may be the same etymon as the preceding), *me(l,n)e ‘tongue’, *mbalaŋ ‘flame’, *mbilaŋ ‘tongue’, *(m,mb)elak ‘light, lightning’ and *mbulikV ‘turn (oneself)’. Reflexes of the following items have more restricted distributions: *kal[a,i]pV ‘casuarina tree’, *walaka ‘testicles’.

*l occurs finally in two etyma: *kindil ‘root’, *saŋgil or *siŋgal ‘hand, finger, claw’

Kalam
Kalam /l/ is a flapped retroflex lateral. *l is kept as l, at least medially and finally.

medial *l > l

*mbalaŋ ‘flame’	maŋlaŋ, malaŋ
*saŋgil ‘hand, finger’	(?) saglaŋ ‘little finger’
*walaka ‘testicles’	walak

final *l > l

*kindil ‘root’	kdl [kɨndɨl]

Apalɨ
final *l >zero

*kindil ‘root’	hɨndɨlɨ (W) (Z. gives gʉndrʉ)

Selepet
Selepet retains *l medially, generally as *l*.

medial *l > l

*mbalaŋ ‘flame’	balam
*(m,mb)elak ‘light, lightning’	belek
*mbilaŋ ‘tongue’	ni-bilam

medial *l > r

*mbulikV ‘turn (oneself)’	purik

Binandere
medial *l > r

*(m,mb)elak ‘light, lightning’	birigi

Telefol

medial *l > l

*kal(a,i)m ‘moon’	*kaliim

An unresolved problem in TNG historical phonology is the origin of the several laterals found in certain languages, chiefly members of the Chimbu-Wahgi group. Some Chimbu-Wahgi languages have three laterals, e.g. an alveolar lateral [l], a velar stop with lateral release [ɬ] or [gl] and a palatalized lateral [l^{y}]. If these are secondary developments in Proto Chimbu-Wahgi, what were the conditioning factors?

Only a small number of Chimbu-Wahgi lexical items containing laterals have possible cognates in other TNG groups. A sample from Middle-Wahgi is given below, speculatively associated with certain eTNG etyma.

eTNG	**Middle Wahgi**
*ma(n,k,l)[a] 'ground'	mał 'ground, soil, world'
*nok 'water'	noł
*ŋaŋ[a] 'baby'	ŋał 'small baby', ŋaŋa 'male child'
*-i(t,l) '2DL verbal suffix'	-ił

Kobon, of the Kalam-Kobon branch of the Madang subgroup, also has three laterals, an alveopalatal, a retroflex flap, and an affricate, but no cognates have been found where they correspond to a lateral in Middle Wahgi.

4.3.2.5 Glides

Widely distributed cognates attesting the glides *w and *y are very scarce. Attribution of the two glides to pTNG rests chiefly on typological arguments. Almost all languages in our sample can be analysed as having two glides, *w* and *y*, occurring word-initially and sometimes finally.

*w occurs only in three etyma, all problematic: *[w]ani 'who?', *walaka 'testicles', *wati 'fence'. The first of these three etyma is quite widely attested but the initial *w is uncertain. The other two are each confined to two major subgroups and there is a strong possibility of borrowing or chance resemblance.

*y is reconstructed initially in *[y]a '3SG free pronoun', *yaka 'bird' and *yara- 'go'. The latter two etyma can be attributed only to stages later than pTNG. *y is reconstructed medially in *kuya 'cassowary' and finally in *mbay 'star'.

Reflexes of etyma with *w and *y are found in just a few languages in our sample.

Chimbu-Wahgi	
initial *y > y	
*ya '3SG'	ye
medial *y is lost	
kuya 'cassowary'	(Kuman kua 'bird')

Kalam
initial *w > w
*walaka ‘testicles’ walak
*wati ‘fence’ wati

initial *w is lost in one dubious comparison
*[w]ani ‘who?’ (?) an

initial *y > y
*yaka ‘bird’ yakt

Asmat
initial *y > zero
*ya ‘3SG’ a

4.3.3 pTNG vowels

4.3.3.1 Vowel systems in the sample languages

Of the ten languages in our sample, four have straightforward five vowel systems of the /a,e,i,o,u/ type, i.e. two front and two back rounded vowels and one low central vowel. The four are Binandere, Kiwai, Middle Wahgi and Telefol. In Telefol each short vowel contrasts with a matching long (or geminate) vowel.

Kalam distinguishes five vowels: /a, e, i, o, u/. It also makes heavy use of predictable epenthetic vowels, often realised as very short [ɨ], but in some contexts as a copy of a neighbouring full vowel. In some cases the epenthetic vowels appear to be, historically, reductions of full vowels.

Apalɨ distinguishes six vowels: /a, e, i, o, u, ɨ/. Wade observes that /e/ and /o/ occur in just a few words and speculates that they derive from recent splitting of /i/ and /u/, respectively.

Kâte and Selepet distinguish six vowels: /a, e, i, ɔ, o, u/.

Asmat has six contrasting vowels: /a, e, ë, i, o, u/, where /ë/ is mid-central. Some of the vowel phonemes have diverse allophones, e.g. the front vowels /i/ and /e/ may each be rounded or unrounded.

Kaeti contrasts seven vowels: /a, e, i, o, ö, u, ü/. /ü/ is high front rounded, /ö/ is mid front rounded.

4.3.2.2 The pTNG vowel system

pTNG probably distinguished at least five vowels, written here as *a, *e, *i, *o, *u. There may have been additional vowels but as yet no clear patterns of correspondences have emerged among the residue of material that does not fit the five vowel hypothesis.

A detailed treatment of the vowels will not be attempted here. As is the case in most language families, TNG languages typically show much irregularity in

their reflexes of reconstructed vowel phonemes, mainly attributable to the familiar processes of analogy, assimilation, apocope, haplology, etc. Thus there are many eTNG etyma where the quality of a reconstructed vowel is indeterminate. When three or more vowels are candidates the vowel is simply represented by *V, otherwise the alternatives are specified, e.g. *kal(a,i)m 'moon', *simb(i,u) 'guts', *takVn[V] 'moon', *mVn[a]- 'be, live, stay', *mVkVm 'cheek', *mo(k,ŋg)Vm 'joint', *nVŋg- 'see, know'.

Table 6 lists the most common reflexes of pTNG vowels in the sample languages. Exceptions are too many to easily tabulate.

Table 8. Most common reflexes of pTNG vowels in the 10 sample languages

pTNG	**a**	**e**	**i**	**o**	**u**
M. Wahgi	a	?	i,e	o	u
Kalam	a	e	i	o	u
Apalɨ (Akɨ)	a	a	i, ɨ	?	u,ɨ
Selepet	a,ɔ	e,o	i	o,ɔ	u,ɔ
Kâte	ɔ,a	e	i	?	u,ɔ
Binandere	a,o	?	i	?	u
Kiwai	a	?	i,e	?	u,o
Telefol	a	?	i	o	u
Kaeti	a,o	?	i	o	u,o,ü
Asmat	a	e	i	?	u

*a is by far the most frequently occurring vowel in pTNG and eTNG etyma, as it is in contemporary TNG languages. The high vowels *i and *u are the next most common. The mid vowels *e and *o are more weakly attested. This frequency ranking is typical of TNG languages with comparable five vowel systems.

Reflexes of pTNG vowels in the 10 sample languages can be found in previous sections, in the examples given for consonant reflexes. The following is a partial list of eTNG etyma given there, arranged by vowels. A few additional examples can be found in section 5.

***a** *am(a,i) 'mother', *ambi 'man', *maŋgV 'compact round object', *mbalaŋ 'flame', *maŋgat[a] 'teeth, mouth', *ŋaŋ[a] 'baby', *mun(a,i,u)ka 'egg', *mbena 'arm', *na '1SG', *na- 'eat', *niman 'louse', *takVn[V] 'moon', *kambu-sumbu 'ashes', and many other sets.

***e** *nde- 'speak', *ke(s,nj)a 'blood', *kend(o,u)p 'fire', *(m,mb)elak 'light, lightning', *me(l,n)e 'tongue', *mbena 'arm', *mbeŋga-masi 'orphan', 'widow and child'.

***i** *imbi ‘name’, *inja ‘tree’, *ambi ‘man, husband’, *kanjipa ‘sun’, *kasin ‘mosquito’ *kasipa ‘to spit’, *kin(i,u)- ‘to sleep’, *niman ‘louse’, *mbilaŋ ‘tongue’, *(ŋg,k)iti-maŋgV ‘eye’, *titi ‘tooth’, *sa(ŋg,k)asiŋ ‘sand’, *simb(i,u) ‘guts’.

***o** *mo(ŋg,k)Vm ‘joint’, *mo ‘penis’, *kend(o,u)p ‘fire’, *k(o,u)ndVC ‘bone’, *k[a,o]ndokV ‘foot, leg’, *k(o,u)ma(n,ŋ)[V] ‘neck, nape’, *n(o,u)man ‘mind, soul’.

***u** *kumV- ‘die’, *k(i,u)tuma ‘night, morning’, *kumbutu ‘wind, breeze’, *kambu-sumbu ‘ashes’, *kuya ‘cassowary’, *m(i,u)ndu ‘nose’, *mundun ‘internal organs’, *mun(a,i,u)ka ‘egg’, *m(o,u)k ‘milk, sap, breast’, *mbulikV ‘turn (oneself)’, *simb(i,u) ‘guts’.

5. RECONSTRUCTING PARTICULAR LEXICAL ITEMS: FIVE CASE STUDIES

When reconstructing lexical forms in a language that existed many thousands of years ago it is to be expected that many reconstructions will show indeterminacies. This section uses several case studies to illustrate some of the issues that arise in reconstructing pTNG forms.

5.1 *kumbutu ‘wind, breeze’

Often one is faced with a formally disparate set of putative cognates, among languages whose phonological histories are obscure. In such cases it is often possible to arrive at an approximate first reconstruction by (a) picking out the longest forms in the cognate set and (b) searching for agreements between widely separated subgroups. The rationale for (a) is that the longer forms are likely to be more conservative and the shorter forms to be the result of erosive changes. The possibility that the longer forms are derived from earlier compounds must be kept in mind (see e.g. 5.2.5). However, if the longer forms are widely distributed across diverse subgroups, this is grounds for concluding that they are conservative, regardless of whether or not they derive from an original compound.

A case in point is the eTNG etymon tentatively reconstructed as *kumbutu ‘wind, breeze’. It can be seen that some of its putative reflexes are trisyllables, others disyllables and yet others monosyllables.

SOUTH-EAST
Koiarian
Barai *uburu*
FINISTERRE–HUON
Uruwa
Sakam *gupi*
Erap
Sauk *gufut*

MADANG
Rai Coast
Usino *kibul*
Mabuso
Sihan *uhe*
Garuh *wus*
N Adelbert
Koguman *(o)gobor*
S Adelbert
Musak *kʉburu*
Sileibi *kunumbu* (metath.)
E Highlands
Chimbu–Wahgi
Narak *kopo*
CENTRAL
Kutubuan
Foe *kuba*
Awyu–Dumut
Pisa *kifi*
Asmat-Kamoro
Kamoro *kimir*

It can be seen that several languages have trisyllabic, or at least triconsonantal roots, which correspond rather closely in form: Barai *uburu*, Musak *kʉburu*, Sileibi *kunumbu*, Sauk *gufut*, Usino *kibul*, Koguman *(o)gobor*, Kamoro *kimir*. These witnesses belong to widely separated subgroups, ranging from the far west of New Guinea (Kamoro) to almost the far east (Koiari). The agreements point to a trisyllable of the form *kumbutu as the most likely ancestral form (on the assumption that *t > r between vowels), with Barai showing loss of *k, Kamoro showing reduction of *mb to *m* and Sileibi showing metathesis of the final two consonants and replacement of *t by *n*: *kumbutu > kutumbu > kunumbu. The other, shorter forms in the set remain as residue, with some as yet unexplained changes. *kifi, kopo, kuba* and *uhe* appear to have lost the final syllable and *wus* the initial syllable and final vowel. In the case of *kumbutu there is no good reason to think that the etymon was a compound.

5.2 *iman and *niman: the *satəm and *centum of pTNG?

Terms for 'louse' are among the most persistent lexical items in many of the world's language families. TNG is no exception. Two obviously related forms for 'louse', *iman and *niman, are widely attested. The following is a sample of more than 100 attested reflexes. The distribution of reflexes of these forms across subgroups presents a historical puzzle.

SOUTH-EAST
Kwalean
Kwale *nomone*
Koiarian
Managalasi *uma*
Aomie *ume*
Binanderean
Suena *dʒimi*
FINISTERRE-HUON
E. Huon
Sene *ime*
Kâte *imeŋ*
Mape *imaŋ*
W. Huon
Sialum *yaman*
Kumukio *imɛn*
Selepet *imen*
Wantoat
Irumu *imɛn*
Warup
Degenan *imeŋ*
Asat *yumun*
Gusap-Mot
Gira *im*
Neko *imiŋ*
MADANG
Except for certain S. Adelbert Range languages, all Madang languages lack initial **n*
Kalam-Kobon
Kalam *iman [yiman]*
Rai
Pulabu *ima*
Tauya *im*
S. Adelbert
Katiati *ñima*
Pondoma *nɯmaŋ*
Faita *ima*
Emerum *iman*
E. HIGHLANDS
Kainantu
Gadsup *numi*
Tairora *numɛ*

Gorokan
Gende *(ti)nima*
Siane *nema*
Fore *numaa*
Chimbu-Wahgi
Tabare *niman*
Kuman *numan*
Maring *numa*
Engan
Enga *lema*
Ipili *lemo*
Kewa *lema*
Wiru
Wiru *nomo*
CENTRAL
Ok
Rather than *[n]iman most of the Ok forms cited here may reflect another etymon with initial *t, possibly meaning 'flea'.
Faiwol *kim*
Telefol *tim*
Lowland Ok
Kati N. *tim*
Kati M. *im*
Asmat-Kamoro
Kamoro *namo*
SOUTH-CENTRAL
Kiwai
Kiwai *nimo*
Gogodala-Suki
Gogodala *ami*
Kolopom
Kimaghama *nome*
Kayagar
Kaugat *numu*
NORTH-WEST
In our sample, most languages in the West region, along with some Timor-Pantar witnesses, reflect *amin rather than *iman.
Kaure
Kosare *mi*
WEST
Mek
Kosarek *ami*
Yale *ami*

Uhunduni	
Uhunduni	*e:mon*
Wissel Lakes	
Ekagi	*yame*
Moni	*amu*
TIMOR	
Oirata	*amin*
Lamma (Pantar)	*hamiŋ*

One might speculate that the variation between *iman and *niman has subgrouping value, e.g. either accretion of *n (e.g. from a fossilised possessive pronoun prefix) or loss of *n might be an innovation marking a subgroup. However, counting against this notion is the fact that if we draw isoglosses around languages that, respectively, have initial *n or lack it, we get groupings that show little or no geographic coherence and which conflict with other subgrouping evidence. In this respect the isoglosses for *iman and *niman resemble those for *satəm and *centum in Indo-European. For example, while *iman reflexes predominate at both the eastern and western extremes of New Guinea, pockets of *niman forms are present at both extremes (e.g. Kwale *nomone,* Kamoro *namo*). And in the middle territory, reflexes of both forms are found, sometimes even in the same lower order subgroup (e.g. S. Adelbert Range). In these circumstances it seems necessary to attribute both *iman and *niman to pTNG, as doublets or dialectal variants.

Certain other changes to the proto-forms may be significant for subgrouping purposes. Members of the Engan group agree in irregularly replacing *n with *l* in this etymon. It is noteworthy that Gogodala, spoken in the southwest of Papua New Guinea, most small subgroups located in West Papua, and possibly some Alor-Pantar languages, reflect *amin rather than *[n]iman. This may represent (i) independent occurrence of metathesis, *iman > amin, in a number of different western languages, (ii) an innovation of an interstage ancestral just to these western languages, or (iii) retention of pTNG *amin, with*[n]iman being an innovation of an interstage ancestral to many non-western languages.

5.3 *maŋgat[a] 'teeth, mouth'

Consider now a different issue posed by another cognate set, referring to 'teeth' and/or 'inner mouth'. The pTNG form can be rather securely reconstructed as *maŋgat[a] because of agreements between witnesses from widely separated subgroups. For example, it can be seen below that Suroi, Kati M., Bimin, Gogodala, Kiwai, Samo and Awyi concur in reflecting three consonants *m-ŋg-t. There is ample support for supposing that the vowel sequence was either *a-a, or *a-a-a, but the question remains whether the pTNG form had two or three syllables. (The default gloss of forms listed below is 'teeth'.)

SOUTHEAST
Mailu
Domu *maʔa* ‘mouth’
FINISTERRE-HUON
Gusap–Mot (both ‘mouth’)
Gira *ma*
Ufim *maŋgo*
Erap
Munkip *maŋ*
MADANG
Kalam-Kobon
Kalam *meg*
Kobon *meg*
Rai
Dumpu *mek*^{h}
Lemio *meg*
Usino *maga*
Usu *maha*
Suroi *maketiŋ*
S Adelbert
Pondoma *maka*
Musak *ma:ki* ‘mouth’
Yaganon
Saep *maŋga-* ‘mouth’
Yabong *mɔŋgose* ‘mouth’
CENTRAL (all ‘mouth’)
Bosavi
Bosavi *mego:f*
E Strickland
Samo *magara*
Bibo *maga:r*
Kubo *moga*
Ok
Kati M *moŋgot*
Bimin *maŋkat-[kun]* ‘chin’
Faiwol *makat-[kalim]* ‘whiskers’
Marind
Marind *mangat*
Awyu–Dumut
Wambon *maŋgot*
SOUTH-CENTRAL
Kiwai
Kiwai *magata*

Turituri *magota*
Gogodala-Suki
Gogodala *magata* ‘teeth, mouth, jaw’
NORTH-CENTRAL
Tami (membership in TNG dubious)
Waris *meŋk*
Awyi *miŋgir*

In favour of the three-syllable hypothesis is the fact that in several languages of different subgroups the reflex is a trisyllable with final *-a*. However, these languages are all located in the Central, South-Central and South-West regions, more or less forming a continuous bloc, and the possibility remains that they belong to a subgroup or an old diffusion area, where final ‘echo’ vowels were added to pTNG disyllables to reinforce a preference for open final syllables. Alternatively, final vowel loss has occurred several times independently in different branches of TNG.

The issue of final vowel loss raises the question of whether stress placement in pTNG or eTNG roots was variable. An argument can be made that it was. It is reasonable to assume that when a language shows final syllable loss, or final consonant loss, the original final syllable was unstressed. In a number of languages certain disyllabic or trisyllabic roots have either been reduced to a single syllable, e.g. *maŋgat[a] has been reduced to *mak, maŋ, meg*, etc., or have lost the coda to the second syllable, yielding *maka, maga, moga*, etc. These reductions are explainable if we assume that the eTNG form was stressed on the penultimate syllable.

However, not all pTNG disyllables and trisyllables are reduced in the daughter languages in question. Thus, although Kalam reduces *maŋgat[a] ‘teeth, mouth’, *simb(i,u) ‘guts’ and *imbi ‘name’ to monosyllabic *meg* [meŋk], *sb* [sɨmp] and *yb* [yí·mp], respectively, it reflects disyllabic etyma such as *maŋgV ‘compact round object’, *(m,mb)elak ‘light, brightness’ and *mbalaŋ ‘flame’ as disyllables *magi* [má·ŋgí·], *melk* [mé·lɨ́k], and *malaŋ* [má·lá·ŋ], respectively, with both syllables taking stress. Such retentions can be explained if we assume that the eTNG etyma in question carried stress on the final syllable. A consideration of the evidence from a range of languages concerning this point would require a separate paper.

5.4 Indeterminacy between a nasal and an obstruent

There are several cases where one subset of cognates reflects a nasal consonant (most often *m or *n) while another subset reflects a matching prenasalised obstruent (most often *mb or *nd) in the same position. Two questions arise here. (1) Is one direction of change more likely than the other? (2) Do the isoglosses have any subgrouping significance?

The following is an example of a cognate set which yields a reconstruction *(m,mb)elak ‘light, lightning, to flash, be bright, etc.’ which shows an

indeterminacy between initial *m and *mb. Independent of the initial consonant reflex, the cognates are divided below into two subsets on semantic grounds. Members of subset (a) refer exclusively to lightning or the flashing of lightning, and members of (b) refer to brightness or light in general, including from sun or fire.

The distributions of *m and *mb reflexes do not follow a neat geographical pattern. It can be seen that both types are present in the Finisterre-Huon and Madang groups and even within lower-order branches of each. Nor do they correlate with the semantic differences between sets (a) and (b).

(m,mb)elak

(a) 'lightning, lightning flash' (N.), or 'to flash, lighten' (V.)

The gloss for forms below should be read as referring to a noun 'lightning, lightning flash' unless otherwise indicated by use of '(V.)', when the gloss should be 'to flash (of lightning)'.

SOUTH-EAST

Binandere

Binandere	*biriga*
Korafe	*biria*

FINISTERRE-HUON

W Huon

Selepet	*belek*

MADANG

Rai Coast

Arawum	*meleŋe-*
Duduela	*amili (fie-)* (V.)
Jilim	*bilen-*

S Adelbert

Moresada	*mera-(tangu-)* (V.)

Croisilles

Kare	*pililia*
Bagup	*(amen) pipile-* (V.)
Bau	*peri (flay-)* (V.)
Sihan	*amera (flay-)* (V.)
Mugil	*meulik (em-)* (V.)
Waskia	*bilik(ma)*
Ukuriguma	*bilika*
Hinihon	*melelek (ewi-)* (V.)
Parawen	*milik (ei-)* (V.)

SOUTH-WEST

Asmat	*mer*

TIMOR

Blagar	*merax*

(b) 'light, brightness' (N.), 'be light (as of fire or sun) (V.)'
Glosses should be read as referring to a noun 'light, brightness' unless otherwise indicated.

FINISTERRE-HUON

E Huon

Kâte	*bɔri?* (V.) 'glitter, flash of lightning, etc.'
	bɔbɔri? (N.) 'lightning, brightness'

MADANG

Kalam-Kobon

Kalam	*melk* (N.) 'light (of day, etc.)', *melk g-* (V.) 'be light'

Rai

Suroi	*buru*
Bongu	*burug*
Arawum	*mele*

S Adelbert

Moresada	*merak*
Sileibi	*(au)mira*
Osum	*mira-* (V.)
Paynamar	*mira*
Faita	*(ni)mera*
Faita	*(ni)mera*

Across TNG, both the weakening of a prenasalised stop to a nasal and the strengthening of a nasal to a prenasalised stop are common irregular sound changes. Thus, it is necessary to build an indeterminacy *(m,mb) into the pTNG reconstruction. However, there are some grounds for favouring *melak as the older form. First, *m*-initial reflexes are somewhat more widespread, occurring both in the east (in three major subgroups of Madang), in the Southwest (Asmat) and in a single witness from the Timor-Alor-Pantar region (Blagar). b-initial reflexes occur in the Southeast (Binandere), Finisterre-Huon and Madang but are not attested in west New Guinea or Timor.

A complicating factor is the likelihood of interference from sound symbolic associations. Suter (p.c.) points out that the sequences *p-l* and *b-l* appear to be sound symbolic for flashing, e.g. Proto Malayo-Polynesian **bilak* 'lightning', German *Blitz*, Latin *fulgur*, English *flash*. Some languages show a proliferation of look-alike terms in this semantic domain, e.g. Ono (Huon Peninsula) *gbilap* 'lightning', *mapalak* 'twinkle', *walatak* 'flash'. Some of the look-alikes in the *(m,mb)elak cognate set may be independent developments or reshaped by sound symbolism. *p-l*, *b-l* sound sound symbolism would favour the independent change of m-initial to b-initial forms rather than the converse.

5.5 Fusion of compounds

The last example concerns terms for 'ashes'. Among cases where sound change has obscured the morphological composition of TNG etyma, this is one of my favourites. TNG languages often distinguish terms for hot ashes and cold or white ashes. An initial search among such terms (beginning with those listed in McElhanon and Voorhoeve 1970) yielded a diverse set of forms, including a number that were more or less resemblant but showed many unexplained irregularities, e.g. Binandere *aβetu*, Koiari *utuvu*, Kovai *tep*, Kâte *dzafe*, Magobineng *dzofo,* Momare *dape*, Bedamini *dasubu*, Bibo *dasuf,* Bongu *sum*, Telefol *kutab*, Faiwol *kutub*, Kaeti *tep*, Wampon *kosep* and Moni *timbwo*, all glossed 'ashes'.

A first comparison of these items suggested a formal reconstruction along the lines of *kV(s,t)V(mb,p)u or *ndVsV(mb,p)u, leaving many indeterminacies. Things became clearer when it was observed that two South Adelbert Range languages and two Rai Coast languages have terms for 'ashes' that appear to reflect an earlier compound: Wadigamam *gawu-sup,* Ikundun *obu-tɨpa*, Urigina *kumbi-sum*, Usino *kumsa-sob*. These point to eTNG *kambu-sumbu, consisting of two nouns which are independently attested as isolable roots: *kambu, probably meaning 'embers, hot ashes', plus *sumbu 'ashes, white ash'. (TNG languages commonly combine two nouns with specific meanings to form a compound with more general meaning, e.g. woman + man 'person, people', girl + boy 'child, children').

In some reflexes of *kambu-sumbu the four-syllable compound has been more or less severely eroded, with one, two, even three syllables lost. However, it appears that not all members of the original set of putative cognates actually reflect a compound. Some languages reflect only *kambu and others reflect only *sumbu. Still others reflect a compound in which one of these morphemes combines with a different element. Voorhoeve (2005:157) points out that Mountain Ok languages reflect a Proto Mountain Ok (pMO) compound that he reconstructs as *uku-tüb or *uku-tëb. Reflexes of *uku meaning 'ashes or fireplace' occur in a number of languages in the western part of Papua New Guinea and West Papua, e.g. Ekagi *ugu* 'fireplace', Telefomin, Mianmin *(u)uk* 'ashes', Gogodala *uku-ru* 'ashes'. Voorhoeve does not attribute a meaning to pMO *tüp or tëp but very likely it continues pTNG *sumbu 'fire, hot ashes'. Whether pMO *uku-tüb was a compound innovated by the Ok subgroup or continued an earlier compound *uku-sumbu is uncertain. Voorhoeve regards the corresponding Dumut form *kosep* as most likely a loan from Ok, because there is good evidence for reconstructing a different Proto Awyu-Dumut etymon for 'ashes', namely *sin(a,o)-kwa(t).

The following is a selection of TNG forms that reflect either the compound *kambu-sumbu or one of its parts. Glosses in contemporary languages mean 'ashes' unless otherwise indicated. Hyphens represent etymological morpheme boundaries, not necessarily recognised as boundaries by speakers of contemporary languages.

***kambu-sumbu 'ashes'**
SOUTH-EAST
Binanderean
Binandere *aβe-tu* (< *kambu-sumbu)
Koiarian
Koiari *u-tuvu* (?) (< *kambu-sumbu)
Mt. Koiari *u-ti* (?) (< *kambu-sumbu)
FINISTERRE-HUON
(all reflecting *sumbu)
E. Huon
Kâte *dzafe*
Magobineng *dzofo*
Ono *dzefe*
Amugen *dzepe* (Amugen is a dialect of Ono)
MADANG
Rai
Usino *kum(sa)-sob* (< *kambu-sumbu)
Urigina *kumbi-sum* (< *kambu-sumbu)
Danaru *kobu(g)* (< *kambu)
N. Adelbert
Pay *tawu(na)* (?) (< *sumbu)
Pila *abu(r)* (?) (< *kambu)
Saki *tawu(r)* (?) < *sumbu)
Tani *tabu(r)* (?) < *sumbu)
S. Adelbert
(all reflecting *kambu-sumbu)
Wadaginam *ga:wu-sup*
Pondoma *o:wu-s* 'fire'
Ikundun *obu-tɨpa* 'fire'
Moresada *uwi-sap* 'fire'
CENTRAL
The Bosavi and E. Strickland languages reflect a compound. The second part continues eTNG * *sumbu* but the etymology of the first part is uncertain.
Bosavi
Bedamini *(da-)subu*
E. Strickland
Samo *(da-)subu*
Bibo *(da-)suf*

The Ok and Awyu-Dumut languages reflect a compound in which the second part derives from *sumbu. In place of *kambu, these languages have *ku-* or *ko-*, which Voorhoeve (2005) derives from *uku 'fire, fireplace'.

Mountain Ok
pMO **ku-tëb* (Healey 1964), PMO **cib* 'white ash'

Bimin (*ku-)teeb*
Faiwol (*ku-)tub*
Telefol (*ku-)tab*
Awyu-Dumut
Wambon (*ko-)sep*
Kaeti (*ko-)tep*
SOUTH-CENTRAL
Kiwai
Kiwai *tuwo* (< *sumbu)
WEST
Wissel Lakes
Moni *timbwo* (< *sumbu)

Shared irregular formal changes may provide evidence for subgrouping. Bedamini and the East Strickland languages share an apparent innovation: irregular replacement of *kambu- by a non-cognate element *nda-. The Ok and Awyu-Dumut reflexes exhibit two irregular changes: (1) reduction of *kambu- to *ko-* or *ku-*, and (2) loss of the final vowel in *sumbu, suggesting that Ok and Awyu-Dumut share a common ancestral form, *ku-sub.

6. CONCLUSIONS: PROBLEMS, PROGRESS, PROSPECTS

It is time to return to the four questions asked at the outset and summarise our conclusions.

(1) Is TNG a valid family? Are there trustworthy criteria for determining membership in TNG?

The answer is clearly yes. The criteria diagnostic for determining membership are agreements in (i) personal pronouns, (ii) other basic vocabulary items, especially those representing some of the 50 or so concepts whose forms are known to have exceptionally high retention rates across language families, (iii) in morphology. Certain groups of languages, e.g. Binanderean, Chimbu-Wahgi, Kainantu-Goroka, Madang, Ok, Awyu-Dumut and Asmat-Kamoro, show enough agreements in (i) and (ii) to establish, beyond reasonable doubt, that they stem from a common ancestor. In some cases the agreements extend to cognate verbal morphology.

(2) Can a clear-cut determination be made for every language?

The answer here is clearly no, both in principle and in practice. In principle, one cannot disprove the claim that any two languages are related. The assertion that languages X and Y are unrelated is only an assertion that no evidence, or no good evidence, has been found that indicates common origin. With regard to quantity and quality of evidence, we can predict, on logical grounds, that the precise limits to membership of the Trans New Guinea family will remain uncertain. There are two reasons for this. First, it is possible, even probable, that some languages in the New Guinea area are very remotely related to Core TNG

languages but retain only the faintest traces of common origin with the latter. Second, it is possible, even probable, that there are some languages which share an immediate common ancestor with Core TNG groups but whose claims to membership in TNG will never be established with certainty because the traces of common ancestry they retain are too fragmentary.

(3) Can we determine whether a particular reconstructed item should be attributed to pTNG or to a later stage? That is to say, is the high order subgrouping of TNG well established?

Question (3) consists of two related but separate questions. The answer to the second is no. The initial dispersal of TNG probably occurred so long ago – perhaps 7 to 10 millennia ago – that little evidence of the sequence of early splits remains. The highest-order branchings within TNG have not been established and quite likely will never be. In any case, it must be assumed that in TNG, as in other language families, differentiation usually took the form of the gradual disintegration of dialect networks rather than sharp splits. We may have to be content with a TNG family tree in which many lower-order branches can be identified, along with a few middle-order branches, but with no clear indications of the primary branchings.

With respect to the first question, the answer is sometimes yes, more often no. The uncertainty about high-order subgrouping means that a cognate set can be attributed to a pTNG etymon with some confidence only if it is represented in low-order subgroups that are geographically widely separated. The more numerous the subgroups and the wider their geographic spread, the higher the level of confidence. At or close to the high end of the scale, for example, are cognate sets for a considerable number of core basic vocabulary items: for most of the personal pronouns and for reflexes of *na- ‘eat’, *kumV- ‘die’, *amu ‘breast’, *am(a,i) ‘mother’, *maŋgat[a] ‘teeth, mouth’, *iti ‘hair’, *ka(nd,t)(e,i)kV ‘ear’, *mb(i,u)t(i,u)C ‘fingernail’, *me(l,n)e ‘tongue’, *kV(mb,p)(i,u)t(i,u) ‘head’, *k(a,o)ndok[V] ‘leg’, *(ŋg,k)a(nd,t)apu ‘skin’, *m(i,u)ndu ‘nose’, *[n]iman ‘louse’, *kasin ‘mosquito, *imbi ‘name’, *inda ‘tree’, *kumbutu ‘wind, breeze’, *takVn[V] ‘moon’ and *maŋgV ‘compact round object’, *(ŋg,k)iti-maŋgV ‘eye’. In almost every case the most widespread cognate sets belong to the set of basic vocabulary items that Cognate sets with more restricted distributions can be attributed, at best, only to an early stage (eTNG). The latter situation is roughly parallel to cases in Indo-European (IE) where a cognate set is restricted to two or three geographically close major subgroups, occurring, say, only in Slavic and Germanic, or Italic and Celtic, or Greek and Indo-Iranian. Indeed, if Hittite and Tocharian are considered to be coordinate with the rest of IE, call the latter ‘Nuclear IE’, then cognate sets confined to Nuclear IE, strictly speaking, do not justify a pIE attribution.

(4) Are the available cognate sets sufficient in number and quality, and widely enough distributed, to arrive at accurate reconstructions of pTNG phonology, lexicon and morphology?

This question is badly formulated in that it asks for a yes or no answer when we are dealing with a scale. The late 1970s and the 1980s saw considerable pessimism about the possibility of making even half-way reliable reconstructions of pTNG or eTNG phonology, lexicon and morphology by applying the Comparative Method from the top down. However, while the pessimists had pointed to various impediments to top-down reconstruction they themselves had hardly tested the waters. When systematic attempts at top-down reconstruction were made in the latter half of the 1990s, quite substantial progress was soon achieved.

Of the 188 eTNG lexical reconstructions examined here, 100 have widely distributed reflexes, represented in subgroups of both the eastern and western halves of New Guinea and these can be attributed at least to a very early stage of TNG. For present purposes the dividing line between eastern and western New Guinea approximates the border between Papua New Guinea and West Papua. Of the remaining lexical reconstructions, 84 have reflexes in subgroups confined to the eastern half of New Guinea. Just four are attested only in the western half of New Guinea. This bias probably reflects two factors: (i) there are considerably more TNG languages in the eastern half of New Guinea than in the western half, (ii) I have searched more diligently among eastern languages than among western languages. Among the lexical reconstructions some are much stronger than others. It is very likely that some of the weaker comparisons will turn out to be spurious.

When it comes to phonological and lexical reconstruction, a language family such as TNG, where only a small residue of well-attested cognate sets survives and where only a small proportion of contemporary languages are well-described, presents special challenges. In such cases, trying to apply the Comparative Method in the usual way, i.e. by seeking to establish sound correspondences between the living languages, will generally yield very meagre results.

In the present paper, four sorts of strategies are used in order to arrive at first very approximate reconstructions of pTNG segmental phonology and lexical forms. The first is to compare the phonological systems of a representative sample of TNG languages with the object of setting up an overarching typology or template inventory of sounds and phonotactic constraints, one that is consistent with the range of observed types. This amounts to a very rough, provisional proto-phonology. The second strategy is to compare possibly related forms (putative cognates) in order to arrive at preliminary lexical reconstructions that are consistent with the provisional proto-phonology. The third step is to try to find regular reflexes of proto-phonemes in a sample of contemporary languages. The fourth is to keep revising the provisional reconstructions as more is discovered.

I concede that the evidence used here to reconstruct pTNG segmental phonology has an eastern bias. Nevertheless, a good part of the segmental phonology can be reconstructed with some confidence. A contrast between three nasal consonants, *m, *n and *ŋ, is well attested, and reflexes of these have been phonetically very stable in daughter languages. A strong case can be made for a

contrast between two series of obstruents, oral vs. prenasalised. Four oral obstruents, *p, *t, *s and *k are well supported. It is uncertain whether [t] and flapped [ř] were positional allophones or whether they contrasted. The former is assumed for the time being. In the prenasalised series a three way contrast between *mb, *nd and *ŋg is strongly indicated, with some evidence for a contrast between *nd and a prenasalised fricative or affricate, *nj. At least one lateral, *l, is reconstructable. Although cognate sets supporting them are few, it is likely that pTNG has two glides *y and *w, because such glides are present in most TNG languages in word-initial and word-final position.

Five vowels are reconstructable: low central *a, plus two front unrounded (*e, *i) and two back rounded (*o, *u) vowels. *a is by far the most common vowel in eTNG etyma, as it is in contemporary TNG languages, followed by *i and *u, with *e and *o least common.

The phonotactic structure of pTNG roots was quite severely constrained. Syllables had the form (C)V, or, root-finally, (C)V(C). Word-internal consonant clusters and vowel clusters were probably not allowed. Prenasalised obstruents probably did not occur in word-final position.

However, in the case of TNG there is only so far one can go using a top-down reconstructive strategy. Most of the lexical reconstructions made so far contain indeterminacies with regard to one or more segments, reflecting unexplained variations in the sound correspondences exhibited by the cognate sets. We must allow that, for the sorts of reasons that Greenberg mentions in connection with French-English sound correspondences (see 4.2.1), many of the indeterminacies in pTNG reconstructions may never be resolved. To have any chance of gaining a more complete and more fine-grained understanding of developments in each lower-order subgroup, it will be necessary to do bottom-up research, or more precisely, to combine top-down with bottom-up work. This in turn will require some of the large gaps in the descriptive record to be filled.

The pool of researchers in Trans New Guinea historical linguistics is small, and there are many subgroups, so we cannot expect rapid progress on a broad front in this domain. However, it is encouraging to see excellent recent or ongoing bottom-up reconstructive work on several TNG subgroups, such as Dutton (2010) on Koiarian, Smallhorn (2011) on Binanderean, Daniels (2006, 2010) on Sogeram, Suter (t.a.) on Huon Peninsula, Voorhoeve (2001) on Awyu-Dumut and Loughnane and Fedden (2011) on the relationship between Oksapmin and the Ok group.

REFERENCES

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.

BROMLEY, MYRON. 1967. The relationships of Grand Valley Dani: a lexicostatistical classification. *Oceania* 37:286-308.

CAPELL, ARTHUR C. 1948. Distribution of languages in the Central Highlands, New Guinea. *Oceania* 19:104-129; 234-253; 349-377.

CHOWNING, ANN. 1987. The supposed Austronesian content of the Trans-New Guinea Phylum. In *A world of language: Papers presented to Professor S.A. Wurm on his 65th birthday*, ed. by D.C. Laycock and Werner Winter, 119-125. Canberra: Pacific Linguistics.

COMRIE, BERNARD. 1986. Haruai verb structure and language classification in the Upper Yuat. *Languages and Linguistics in Melanesia* 17:140-160.

COMRIE, BERNARD. 1989. Genetic classification, contact, and variation. In *Synchronic and diachronic approaches to linguistic variation and language change. Georgetown University Round Table on Languages and Linguistics 1988*, ed. by T.J. Walsh, 81-93. Washington DC: Georgetown University Press.

DANIELS, DON R. 2006. Proto Sogeramic. Comparative reconstruction in central Madang Province, Papua New Guinea. Honors thesis. Dartmouth College.

DANIELS, DON R. 2010. A preliminary phonological history of the Sogeram languages of Papua New Guinea. *Oceanic Linguistics* 49(1): 163-193.

DENHAM, Tim. 2005. Agricultural origins and the emergence of rectilinear ditch networks in the highlands of New Guinea. In Pawley, Attenborough *et al.*, 329-361

DUTTON, THOMAS E. 2010. *Reconstructing proto Koiarian: the history of a Papuan language family.* Canberra: Pacific Linguistics.

DYEN, ISIDORE, A.T. JAMES AND J.W.L. COLE. 1967. Language divergence and the estimated word retention rate. *Language* 43(1):150-171.

FLIERL, WILHELM AND HERMANN STRAUSS. 1977. *Kâte dictionary*. Canberra: Pacific Linguistics.

FOLEY, WILLIAM A., 1986. *The Papuan languages of New Guinea.* Cambridge: Camnbridge University Press.

FOLEY, WILLIAM. 2010. Language contact in the New Guinea region. In *The handbook of language contact*, ed. by Raymond Hickey, 793-810. Oxford: Wiley-Blackwell.

GREGERSON, KENNETH AND MARGARET HARTZLER. 1987. Towards a reconstruction of Proto-Tabla-Sentani phonology. *Oceanic Linguistics 26* (1-2): 1-29.

GREENBERG, JOSEPH H. 1971. The Indo–Pacific hypothesis. In *Current trends in linguistics* vol. 8*: Linguistics in Oceania,* ed. by Thomas A. Sebeok, 807-871. The Hague: Mouton.

GREENBERG, JOSEPH H. 2005. *Genetic linguistics. Essays on theory and method,* ed. by William Croft. New York: Oxord University Press.

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, Research School of Pacific Studies, Australian National University.

HEALEY, ALAN. 1970. Proto-Awyu-Dumut phonology. In *Pacific Linguistic studies in honour of Arthur Capell*, ed. by S.A. Wurm and D.C. Laycock, 997-1063. Canberra: Pacific Linguistics.

HEALEY, PHYLLIS AND ALAN HEALEY. 1977. *Telefomin dictionary.* Canberra: Pacific Linguistics.

HEESCHEN, VOLKER. 1978. Review of Wurm (ed.) 1975. *Anthropos* 72:978-980.

HOLMAN, ERIC W., SOEREN WICHMANN, CECIL H. BROWN, VIVEKA VELUPILLAI, ANDRE MUELLER AND DIK BAKKER. 2009. Explorations in automated language classification. *Folia Linguistica* 42/ 3-4: 331-354.

HOLMAN, ERIC W. , SØREN WICHMANN, CECIL H. BROWN, VIVEKA VELUPILLAI, ANDRE MÜLLER AND DIK BAKKER. 2009. Explorations in automated language classification. *Folia Linguistica* 42/ 3-4: 331-354.

KING, COPLAND. 1927. *Grammar and dictionary of the Binandere language, Mamba River, North Division, Papua.* Sydney: D.S. Ford.

LANG, RANIER. 1975. Review of Wurm (ed.) 1975. *Kivung, Journal of the Linguistic Society of Papua New Guinea* 9/1: 72-80.

LAYCOCK, DONALD C. AND WERNER WINTER (eds). 1987. *A world of language. Papers presented to Professor S.A. Wurm on his 65th birthday.* Canberra: Pacific Linguistics.

LOUGHNANE, ROBYN AND SEBASTIAN FEDDEN. 2011. Is Oksapmin Ok? A study of the genetic relationship between Oksapmin and the Ok Language. *Australian Journal of Linguistics* 31/1: 1-42.

MCELHANON, KENNETH A. 1967. Preliminary observations on Huon Pensinsula languages. *Oceanic Linguistics* 6: 1-45.

MCELHANON, KENNETH A. 1970. Lexicostatistics and the classification of the Huon Peninsula languages. *Oceania* 40(3):214-231.

MCELHANON, KENNETH A. 1975. North-Eastern Trans–New Guinea Phylum languages. In Wurm (ed.) 527-567.

MCELHANON, KENNETH A. and N.A. MCELHANON. 1970. *Selepet dictionary.* Canberra: Pacific Linguistics.

MCELHANON, KENNETH A. AND C.L. VOORHOEVE. 1970. *The Trans–New Guinea Phylum: explorations in deep-level genetic relationships.* Canberra: Pacific Linguistics.

PAWLEY, ANDREW. 1995. C.L. Voorhoeve and the Trans New Guinea Phylum hypothesis. In Baak et al.,1995, 83-122.

PAWLEY, ANDREW. 1998. The Trans New Guinea Phylum hypothesis: a reassessment. In *Perspectives on the Bird's Head of Irian Jaya, Indonesia. Proceedings of the conference, Leiden, October 13-17, 1997,* ed. by Jelle Miedema, Cecilia Ode and Rien A.C. Dam, 655-690. Amsterdam/Atlanta: Rodopi B.V.

PAWLEY, ANDREW. 2000. Notes on early Trans New Guinea verb morphology: reconstructing subject suffixes. Printout. Dept. Linguistics, RSPAS, Australian National University.

PAWLEY, ANDREW. 2001. The Proto Trans New Guinea obstruents: arguments from top-down reconstruction. In Pawley, Ross & Tryon (eds), 261-300.

PAWLEY, ANDREW. 2005. The chequered career of the Trans New Guinea hypothesis. In Pawley, Attenborough *et al.*, 67-107.

PAWLEY, ANDREW. 2009a. The role of the Solomon Islands in the first settlement of Remote Oceania: bringing linguistic evidence to an archaeological debate. In *Austronesian historical linguistics and culture history: a festschrift for Robert Blust,* ed. by Alexander Adelaar and Andrew Pawley, 515-540. Canberra: Pacific Linguistics.

PAWLEY, ANDREW. 2009b. Greenberg's Indo-Pacific hypothesis: an assesssment. In *Discovering history through language: Papers in honour of Malcolm Ross*, ed. by Bethwyn Evans, 153-180. Canberra: Pacific Linguistics.

PAWLEY, ANDREW. 2011. Some Trans New Guinea etymologies. Printout. Dept. Linguistics, College of Asia & the Pacific, Australian National University.

PAWLEY, ANDREW AND RALPH BULMER. 2011. *A dictionary of Kalam with ethnographic notes.* Canberra: Pacific Linguistics.

PAWLEY, ANDREW, ROBERT ATTENBOROUGH, JACK GOLSON AND ROBIN HIDE (eds). 2005. *Papuan pasts: cultural, linguistic and biological histories of Papuan-speaking peoples.* Canberra: Pacific Linguistics.

PAWLEY, ANDREW, MALCOLM ROSS AND DARRELL TRYON (eds). 2001. *The boy from Bundaberg: essays in Melanesian linguistics in honour of Tom Dutton.* Canberra: Pacific Linguistics.

PAWLEY, ANDREW AND HARALD HAMMARSTRÖM. to appear. The Trans New Guinea family. In Bill Palmer ed., *The languages and linguistics of New Guinea: A comprehensive guide.* Berlin: de Gruyter Mouton.

RAMSEY, EVELYN. 1975. *Middle-Wahgi dictionary*. Mt Hagen: Church of the Nazarene.

RHODES, RICHARD A. 1997. On pronominal systems. In *Indo-European, Nostratic and beyond: festschrift for Vitalij V. Shevoroshkin,* ed. by Irén Hegedüs, Peter A. Michalov and Alexander Manaster Ramer, 293-319. Journal of Indo-European Studies monograph 22. Washington D.C.: Institute for the Study of Man.

ROSS, MALCOLM D. 1995. The great Papuan pronoun hunt: recalibrating our sights. In Baak *et al.,* 139-168.

ROSS, MALCOLM D. 1996. Contact-induced change and the comparative method: cases from Papua New Guinea. In *The Comparative Method reviewed: regularity and irregularity in language change,* ed. by Mark Durie and Malcolm Ross, 180-217. New York, OUP.

ROSS, MALCOLM D. 2000. Pronouns as preliminary evidence for grouping Papuan languages. Computer file, Dept of Linguistics, Collegd of Asia & the Pacific,
Australian National University.

ROSS, MALCOLM D. 2001. Is there an East Papuan phylum? Evidence from pronouns. In Pawley *et al.* (eds) 2001, 301-321.

ROSS, MALCOLM D. 2005. Pronouns as a preliminary diagnostic for grouping Papuan languages. In Pawley, Attenborough *et al.*, 15-66.

SHAW, DANIEL. 1983. A tentative classification of the languages of the Mt Bosavi region. In *The linguistic situation in the Gulf District and adjacent areas, Papua New Guinea*, ed. by Karl Franklin ed., 189-215. Canberra: Pacific Linguistics.

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.

SUTER, EDGAR. to appear. A historical-comparative grammar of the Huon Peninsula languages. PhD thesis, University of Cologne.

TADMOR, URI, MARTIN HASPELMATH AND BRADLEY TAYLOR. 2010. Borrowability and the notion of basic vocabulary. *Diachronica* 27(2):226-246.

THURSTON, WILLIAM R. 1987. *Processes of change in the languages of north-western New Britain.* Canberra: Pacific Linguistics.

VOEGELIN, CARL, AND FLORENCE VOEGELIN. 1965. *Languages of the world. Indo-Pacific fascicle five. = Anthropological* Linguistics vol. 5, no. 5. Bloomington: University of Indiana.

VOORHOEVE, CLEMENS L. 1965. *The Flamingo Bay dialect of the Asmat language*. The Hague: H.L. Smits.

VOORHOEVE, CLEMENS 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.

VOORHOEVE, CLEMENS L. 1980. *The Asmat languages of Irian Jaya*. Canberra: Pacific Linguistics.

VOORHOEVE, CLEMENS L. 1987. Worming one's way through New Guinea: the chase of the peripetatic pronouns. In Laycock and Winter, 1987, 709-727.

VOORHOEVE, CLEMENS L. 2001. Proto-Awyu-Dumut phonology II. In Pawley, *et al.* eds, 2001, 361-381. Canberra: Pacific Linguistics.

VOORHOEVE, CLEMENS L. 2005. Asmat-Kamoro, Awyu-Dumut and Ok: an enquiry into their linguistic relationships. In Pawley, Attenborough *et al.*, 2005 145-166.

WADE, MARTHA. n.d. Dictionary of the Apali language. Printout. Madang: Pioneer Bible Translators.

WILSON, DARRYL. 1969. The Binandere language family. Pacific Linguistics Series A-18, 65-86. Canberra: Pacific Linguistics.

WURM, STEPHEN A. 1960. The changing linguistic picture of New Guinea. *Oceania* 31: 121-135.

WURM, STEPHEN A. 1964. Australian New Guinea Highland languages and the distribution of their typological features. *American Anthropologist* 66/4(2):79-97.

WURM, STEPHEN A. 1965. Recent comparative and typological studies in Papuan languages of Australian New Guinea. *Lingua* 15:373-399.
WURM, STEPHEN A. 1971. The Papuan linguistic situation. In *Current trends in linguistics* vol.8. *Oceania*, ed. by T.E. Sebeok, 541-657. The Hague: Mouton.
WURM, STEPHEN A. 1973. The Kiwaian language family. In *The linguistic situation in the Gulf District and adjacent areas, Papua New Guinea*, ed. by Karl Franklin, 217-260. Canberra: Pacific Linguistics.
WURM, STEPHEN A. 1975a. Personal pronouns. In Wurm, ed., 191-217.
WURM, STEPHEN A. 1975b. The application of the comparative method to Papuan languages: general and Highlands. In Wurm, ed., 237-261.
WURM, STEPHEN A. 1982. *Papuan languages of Oceania.* Ars Linguistica 7. Tübingen: Günter Narr.
WURM, STEPHEN A. (ed.) 1975, *New Guinea area languages and language study,* vol.1: *Papuan languages and the New Guinea linguistic scene.* Canberra: Pacific Linguistics.
WURM, STEPHEN A. AND SHIRO HATTORI. 1981-83, *Language atlas of the Pacific area.* Canberra: Australian Academy for the Humanities in collaboration with the Japanese Academy.
WURM, STEPHEN A., C.L. VOORHOEVE AND K. McELHANON. 1975. The Trans-New Guinea Phylum in general. In Wurm, ed., 299-322.
Z'GRAGGEN, JOHN A. 1971. *Classificatory and typological studies in languages of the Madang District.* Canberra: Pacific Linguistics.
Z'GRAGGEN, JOHN A. 1975. The Madang–Adelbert Range sub-phylum. In Wurm ed., 569-612.
Z'GRAGGEN, JOHN A. 1980a. *A comparative word list of the Rai Coast languages, Madang Province, Papua New Guinea.* Canberra: Pacific Linguistics.
Z'GRAGGEN, JOHN A. 1980b. *A comparative word list of the Northern Adelbert Range languages, Madang Province, Papua New Guinea.* Canberra: Pacific Linguistics.
Z'GRAGGEN, JOHN A. 1980c. *A comparative word list of the Mabuso languages, Madang Province, Papua New Guinea.* Canberra: Pacific Linguistics.
Z'GRAGGEN, JOHN A. 1980d. *A comparative word list of the Southern Adelbert Range languages, Madang Province, Papua New Guinea.* Canberra: Pacific Linguistics.