

ON THE HISTORY OF THE TANNA NUMERALS AND NUMBER-MARKERS

John Lynch
(University of Papua New Guinea)

0. INTRODUCTION

In attempting to account for the development of Proto-Oceanic (POC) reconstructed etyma in the languages of Tanna, Southern New Hebrides (Lynch 1978), I was struck by the considerable amount of variation between the expected derivation of the POC numerals and their actual forms in the various languages concerned. I was also impressed by variation between the form of the numerals 'two', 'three', and 'four' and the shape of corresponding grammatical affixes marking dual, trial, and plural. I attempt to account for both types of variation in this paper. In so doing, I will suggest modifications to the form of the existing POC reconstructed numerals, and will also propose that POC may well have had a series of number-marking affixes phonologically distinct from the numerals.¹

0.1 THE DATA

The data on which this paper is based are drawn largely from five speech-traditions on the island of Tanna in the Southern New Hebrides (SNH). The division between language and dialect in Tanna is not an easy one to draw. There is a considerable amount of evidence to suggest that North Tanna (NTN), Lenakel (LEN), and Whitesands (WSN) form a single dialect-chain occupying the northern half of the island; that Nivhaal (NVH) is one member of a dialect-chain occupying a central strip and most of the south-west of the island; and that Kwamera (KWM) is one member of a south and south-east dialect-chain. (See map for exact locations). Nevertheless, for the purposes of this paper I shall speak of the five speech-traditions referred to above as if they were distinct languages.²

The Tanna languages have a quinary numeral system; the numerals in each language are given in Table I. All numerals

MAP

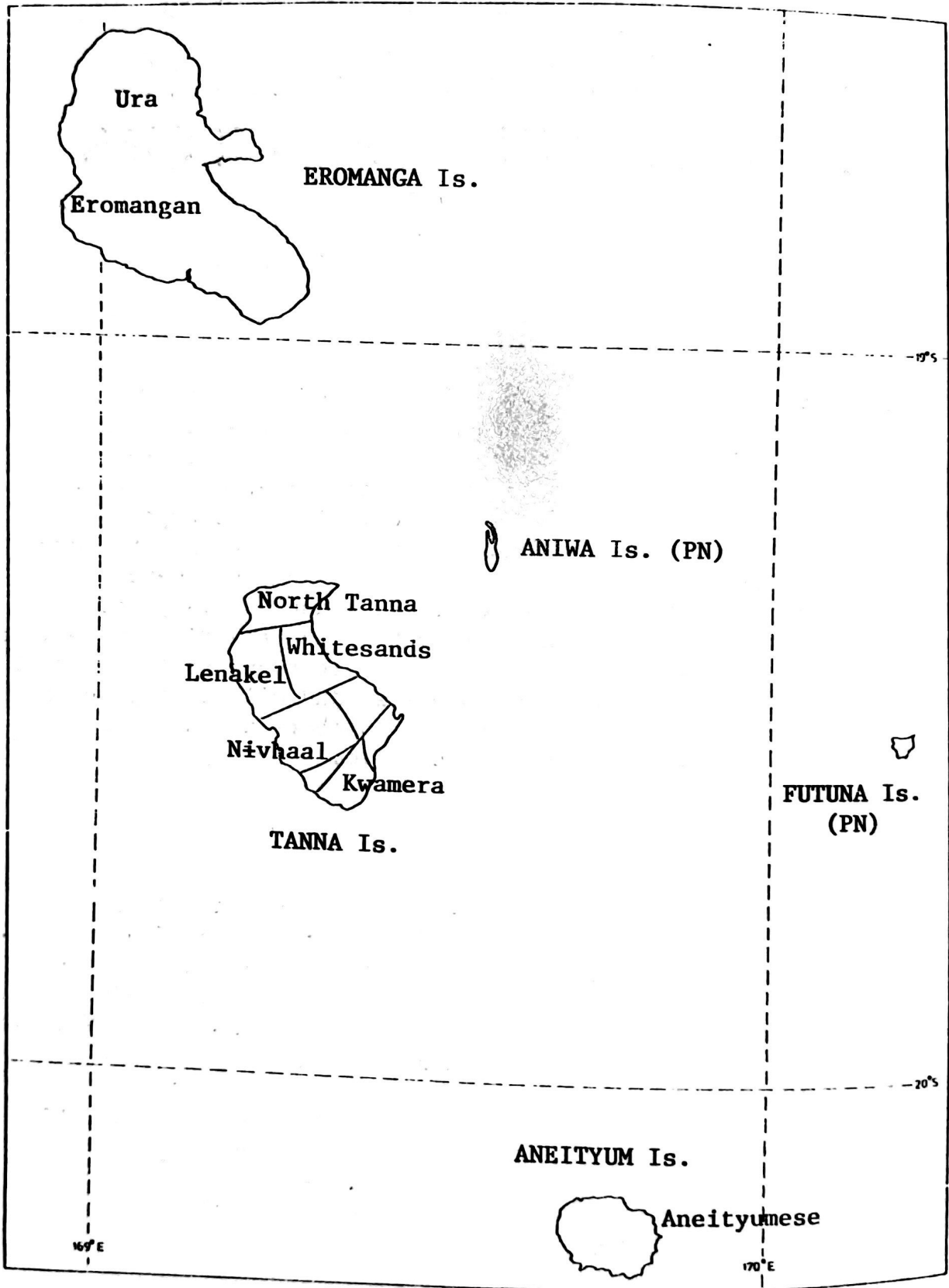


TABLE I					
TANNA NUMERALS					
	'one'	'two'	'three'	'four'	'five'
NTN	kitia	kɪiu	kɪsɪl	kuvɪt	karilɪm
LEN	karena	kiu	kɪsil	kuvɪr	katilum
WSN	katia	kɪiu	kɪsɪl	kuvɪt	karilɪm
NVH	kɪlikiana	kɪlalu	kɪsisɪl	kuas	kɪlkɪlɪp
KWM	rɪtia	kɪru	kahar	kefa	kɪrirɪm

in all of the languages (except (KWM *rɪtia* 'one')) show prefixation of *k(V)* to a root usually derivable from a POC reconstructed numeral. This prefix is probably cognate with Pawley's (1976:11) reconstruction **ka* 'preposed to numerals ... used in counting a sequence'. The presence of this form in Tanna suggests that it should be assigned to POC.³

Although numerals above six are today usually borrowed from Bislama (New Hebrides Pidgin English), in the past they were formed by compounding, using the base 'five'. NTN and LEN merely showed concatenation of the appropriate numerals:

LEN *katilum - katilum - kiu* 'twelve'
 five - five - two

WSN, NVH, and KWM used a ligative *mɪ*, which probably derives from POC **ma*, 'and':

NVH *kɪlkɪlɪp - mɪ - kɪlalu* 'seven'
 five - and - two

The forms for twenty and above included the word for 'man', probably as a result of counting on fingers and toes:

LEN *ieramim karena r - ɪka* 'twenty'
 man one he - is:not

In addition to the numerals, the Tanna languages show two sets of affixes marking grammatical number. One of these sets occurs suffixed to the non-singular pronouns; indeed, these pronouns cannot occur without a suffix marking dual, trial, or plural number. Full pronoun sets in the Tanna languages are given below in Table II, and the form and history of the suffixes will be discussed in the relevant sections below.

The other set of number-marking affixes are prefixes to verbs marking the number of the subject. Note the use of these prefixes in the LEN examples in Table II.

0.3 ORGANISATION

In the first two sections of the paper I will discuss the numerals 'one' and 'five'. I have broken numerical order here since, unlike the other numerals, neither 'one' nor 'five' is associated with any number-marking affix. The next three sections discuss, respectively, the numeral 'two' and the dual affixes, 'three' and the trial, and 'four' and the plural. Implications for POC lexicon and grammar are summarised in section 6.

1. "ONE"

The numerals for 'one', given above in Table I, are recapitulated below:

NTN	kitia
LEN	karena
WSN	katia
NVH	kilikiana
KWM	ritia

LEN and NVH show what is probably a suffix *-na* which has not as yet been identified. All the languages show a reflex of the POC prefix **ka* except KWM: KWM initial *r-* may be the third person singular verbal prefix. However, the KWM form is suspect in other ways: KWM medial *t* does not normally correspond with NTN WSN *t*, LEN *r*, and NVH *l* - the expected KWM consonant is *r* - and so the KWM form may not be cognate at all.

Ignoring the supposed suffix *-na*, the NTN, LEN and WSN forms suggest a PTN reconstruction **t(aou)V*, while the NVH form suggests PTN **t(aou)kVV*. (The vowel of the initial syllable cannot be **i* or **e*, since in this environment **t > s*.)

There are a number of reconstructed forms for 'one' in the Oceanic family. For Proto-Eastern Oceanic (PEO),⁴ Pawley (1972) has reconstructed **(n)sa*, **ta(n)sa*, and **(n)sakai*. Capell (1943) reconstructed Proto-Austronesian (PAN) **ta+kai* on Oceanic evidence, but his evidence for the initial **t* was not convincing, and Grace (1969) accepted as POC only the second syllable, which he reconstructed as **-kai*. From these various reconstructions we might abstract the forms **(n)sa*, **ta*, and **kai*, and suggest that each (either alone or in combination with other forms) has something to do with the meaning 'one'.

In support of his PEO reconstruction **ta(n)sa* 'one', Pawley (1972:52) quotes Proto-Polynesian (PPN) **taha*, Nogugu *tamo*, Baki, Tasiko *tai*, Arosi *ta*, Nggela *eta*, and Vaturanga *tasa*. He also notes Roviana, Lakalai *tasa*, Willaumez (Bula) *tara*, Motu *tamona*,

and also Motu *ta..ta* 'each',⁵ presumably to show that the form might well be reconstructed for POC.

The Nogugu, Baki, Tasiko, Arosi, Nggela, and Motu forms do not reflect the second syllable of his reconstruction. Moreover, Baki and Tasiko show what appears to be an intrusive *i*, which is found again in Tolai *tikai*, in the Siau languages of the Sepik (Ali *tei*, Suain *batai*, Kairiru *tai* (Ross 1977)), and in Sio (Rai Coast) *taitu* (Lincoln 1976:1), as well as in the Tanna forms. I suggest that the PTN form which I reconstructed above as **t(aou)V* is in fact **tai* (with subsequent metathesis of the vowels in four of the languages and reduction of **ai* > *e* in LEN). A POC form **ta(i)* should thus be reconstructed on the basis of the forms discussed above, and PEO **ta(n)sa* should be rewritten as **ta(i) - (n)sa*.

The NTN, LEN, and WSN forms for 'one' thus derive regularly from POC **ka - ta(i)* with metathesis or reduction of the vowels in the second morpheme. The NVH form derives from POC **ka - ta(i) - kai*, with metathesis of the vowels in the third morpheme; **a* in the second morpheme has also been lost, but only after the rule changing **t* > *s* before **i* had ceased to operate.

2. "FIVE"

The numerals for 'five' are given below:

NTN	karilim
LEN	katilum
WSN	karilim
NVH	kilkilip
KWM	kiririm

The NVH form is almost certainly not cognate with either the other Tanna forms or any POC reconstruction.

The numerals in the other four languages appear to derive from a PTN form **ka - nsa - lvm*. The second morpheme must be reconstructed positively as **nsa* and not ambiguously as **nsV* since before **a* POC **ns* > NTN WSN *n*, LEN *t*, whereas elsewhere POC **ns* > NTN LEN WSN *h*.

The root **lvm* obviously derives from POC **lima* 'hand, five'. The morpheme **nsa* is cognate, I believe, with PEO **(n)sa* 'one'; a POC form **nsa* 'one' must thus be reconstructed. The PTN reconstruction **ka - nsa - lima* 'numeral-one-hand, numeral-one-five' is comparable with the use of morphemes meaning 'one' in counting units of ten in many other Oceanic languages (e.g., PEO **(n)sa - ŋa - pulu* 'ten'), and is easily understandable in terms of a five-base counting system. In suggesting that the Tanna

forms (except for NVH) derive from the combination of the morphemes *ka, *nsa, and *lima, the only remaining matter of interest is why the form *nsa 'one' should be used here when the forms *ta(i) and *kai are used in forming the numeral 'one'. I have no answer to this question as yet.

3. "TWO" AND THE DUAL

3.1 THE NUMERAL

The numerals for 'two' are given below:

NTN	k̄iu
LEN	kiu
WSN	k̄iu
NVH	k̄ilalu
KWM	k̄iru

The NVH form shows reduplication, which also appears in 'three' and possibly 'five'. All the Tanna numerals appear to derive regularly from POC *ka - dua with regular loss of the final vowel.⁶

3.2 THE DUAL PRONOUN SUFFIX

The dual pronouns in the Tanna languages are re-presented here in Table V for convenience.

TABLE V				
<u>TANNA DUAL PRONOUNS</u>				
	<i>1st inc.</i> ¹	<i>1st exc.</i>	<i>2nd</i>	<i>3rd</i> ¹
NTN	kihlao	itlao ²	it̄imlao ²	ilao
LEN	kalau	kamlau	kamilau	ilau
WSN	kihlau	it̄imlhau ²	it̄ilau ²	ihlau
NVH	k̄ilau	k̄imlu	k̄imilu	iliu
KWM	k̄irau	k̄imrau	k̄imirau	irau

NOTES. 1. There is a rule of coronal simplification in most Tanna languages (see Lynch 1975) such that, if two alveolar consonants come together across a morpheme-boundary, the first is lost (or, in WSN, replaced by /h/). Thus the underlying LEN forms for the first inclusive and third dual pronouns are /kat-lau/ and /il-lau/ respectively; and similarly for the other languages.

2. The pronominal part of this form does not appear to be derivable from POC.

On the basis of these data and those in Table II, the following dual suffixes to pronouns can be identified:

NTN	-lao
LEN	-lau
WSN	-lau
NVH	-lau, -lu
KWM	-rau

I am unable to explain the variation in NVH between *-lau* in the first inclusive and *-lu* in the other persons (third person *iliu* being regarded as a reduction of ***ili-lu*). This apart, the dual suffixes would derive from a PTN form **-dau*.⁷

Now most other Oceanic languages derive their dual pronouns (if they have any at all) by suffixing **dua* 'two' to their plural pronouns, although some show phonological reduction by dropping the last vowel of the numeral (Pawley 1972:68-75). This last development is not unusual. When two lexical items form a compound which is frequently used, reduction often takes place - e.g., English (some dialects) *forehead* (/fɔrəd/), *cupboard* (/kʌbəd/), and so on. This may well explain the origin of the NVH allomorph *-lu*. Apart from this instance, however, the Tanna languages show not reduction but what appears to be metathesis of the vowels of POC **dua* 'two'.

At least one other Oceanic language also shows this same development. Compare the Bauan Fijian (BAU) plural and dual pronouns:

	<i>plural</i>	<i>dual</i>
1st inc.	keda	kedaru
1st exc.	keimami	keirau
2nd	kemunī	keirau
3rd	ira	irau

Although first inclusive shows *-ra*, first exclusive and third show *-rau*, the third dual form probably being a contraction of an earlier ***ira-rau*. The second person shows *-drau*, but this is probably due to the presence of the nasal in the preceding syllable. The postulated development ***kemunī-rau* > ***kemun-rau* > *kemu-drau* is supported by the second person trial form *kemudou*. A trial suffix *-tou* can be isolated (*kedatou* '1st inc. trial', *keitou* '1st exc. trial', *iratou* '3rd trial'), and the second person trial *kemudou* derives from ***kemunī-tou* in the same way as the second person dual *kemudrau* derives from ***kemunī-rau*.

Thus with the exception of the first inclusive in BAU and most persons in NVH, the Tanna languages and BAU agree in showing a dual pronoun suffix deriving from an ancestral form **-dau*, which appears to be a metathesised form of POC **dua* 'two'. It may be premature at this stage to argue from reasons of

economy by stating that this metathesis occurred only once, in a language ancestral to both BAU and the Tanna languages (i.e., POC or one of its earliest descendants). In any case, I will postpone until section 6 any general discussion of the development of number-affixes in Oceanic.

3.3 THE DUAL NUMBER-PREFIX TO VERBS

Each Tanna language has at least two allomorphs of the morpheme which marks dual subject, although the conditioning differs from language to language.

KWM differs from the other languages in the position in the verb in which the number-prefixes occur: while in the other languages the number-prefix immediately precedes the verb root, KWM number-prefixes precede the tense-prefix (if one is used). In KWM, the dual prefixes are *rou-* and *r-*: *rou-* occurs when no tense-prefix is used (i.e., immediately before the verb root), while *r-* occurs if a tense-prefix is used. One possible explanation of the variation is that *r-* is a contraction of *rou-*, occurring when a series of prefixes is used.

In the other languages, the form of the dual prefix depends on the nature of the initial segment or segments of the root. The distribution of the allomorphs of the dual prefix in these four languages is given below in Table VI.

I would like to view the occurrence in LEN of the dual prefix *u-* before verbs beginning with *ou*, as well as *ua-* before verbs beginning with *ua*, as exceptions to a more general rule that *o-* and *u-* initial verbs take the dual prefix *ia-*. The latter case may be explained as an assimilation: i.e., the sequence **ia-ua* became *ua-ua*. The former case might be explained thus: since verbs with initial *aw* (phonetically [ɔw]) regularly take the dual prefix *u-*, this rule may have been extended to verbs with initial *ou*, which is also phonetically [ɔw].

Whatever the explanation, if these two exceptional cases are ignored, then we can see that two dual allomorphs can be reconstructed (PTN **da-* > NTN LEN WSN *ia-*, NVH *la-*, and PTN **u-* > NTN LEN WSN NVH *u-*); that all languages show a reflex of PTN **da-* before consonant-initial verbs; and that before vowel-initial verbs the following patterns emerge:

NTN & LEN				WSN & NVH			
i	u	*da-		i	u	*da-	
e	ɨ		o	e	ɨ	o	*u-
*u	a			a			

TABLE VI

DISTRIBUTION OF DUAL PREFIXES IN NTN, LEN, WSN, NVH

<i>Verb-initial segment(s)</i>	<i>NTN</i>	<i>LEN²</i>	<i>WSN</i>	<i>NVH²</i>
a	u-	u-	u-	u-
e	u-	u-	u-	u-
ɨ	u-	u-	u-	u-
ou	ia-	u-	u-	u-
oū ¹	ia-	ia-	u-	u-
ua	ia-	ua-	ia-	la-
uā ¹	ia-	ia-	ia-	la-
i	ia-	ia-	ia-	la-
consonant ²	ia-	ia-	ia-	la-

NOTES. 1. The notation \bar{u} (or \bar{a}) means any segment except *u* (or *a*).

2. In LEN and NVH, verbs beginning with /h/ take the prefix appropriate to the following vowel - i.e., initial /h/ is ignored in assigning the dual prefix, and thus verbs beginning with /ha/ take the prefix *u-* while verbs beginning with /hi/ take the prefix *ia-*, NVH *la-*. The reason for this is that there has been a general forward movement of /h/ in these two languages, such that inherited sequences of the form **(C)h* have become *hV*. Presumably, the rules assigning dual prefixes came into operation before the rule whereby /h/ began to apply.

In WSN and NVH, *i* and *u* become glides when adjacent to another vowel. These two languages thus have the general rule that the dual prefix is a reflex of **da-* before a consonant or a "potential consonant" (i.e., *i* and *u*), but a reflex of **u-* before a non-consonant.

NTN also has the general Tanna characteristic that *i* and *u* become glides when adjacent to vowels. It is the only Tanna language, however, where *o* also becomes a glide in this environment. Thus NTN has the same general rule as WSN and NVH: **da-* before a consonant or "potential consonant", **u-* before a non-consonant.

LEN remains exceptional, since *o* never becomes a glide. It *might* be the case that, since LEN and NTN are members of the same dialect-chain, LEN once had the NTN rule whereby *o* became a glide, but later lost this rule after the morphophonemics of the dual prefix had been established. The fact that WSN is also a member

of the same dialect-chain, however, - and is probably more closely related to NTN than LEN is - suggests that this hypothesis is a weak one. We may be forced to accept the less attractive view that LEN borrowed the NTN rule of treating *o* as a "potential consonant" for the purposes of the assignment of the dual prefix.

With one or two minor exceptional cases in LEN, I think it safe to assume that the PTN dual prefix had two allomorphs: **da-* before consonants and "potential consonants", and **u-* before non-consonants. When taken together with the KWM dual prefix *rou-*, these forms suggest a pre-PTN dual prefix **dau-* (i.e., a morpheme identical in form to the dual suffix to pronouns). In consonant-initial verbs, **dau-* has reduced to **da-*, probably as a result of consonant cluster simplification: **dau-* would always end in the consonant [w], and when prefixed to a verb beginning with a consonant the [w] could easily become redundant (just as, in LEN, the verb *sɨm* 'to come from' becomes *sɨ* when followed by a consonant-initial directional suffix). Thus the derivation of LEN *ia-kɨn* 'dual-eat' would have proceeded roughly as follows:

<i>*dau-kani</i>	PRE-PTN FORM
<i>daw-kani</i>	PHONETIC EQUIVALENT
<i>da -kani</i>	REDUCTION
<i>ia -kɨn</i>	OTHER REGULAR SOUND CHANGES

Before vowel-initial verbs, however, there would be no phonetic rationale for the loss of [w]. But it is reasonable to assume, I think, that the initial CV of the prefix could be lost without creating ambiguities, since there are no other verbal prefixes of the form *u-*. Thus I would suggest that the derivation of a verb like LEN *u-arɨk* 'dual-live' would have proceeded roughly as follows:

<i>*dau-a-toka</i>	POC/PTN FORM ⁸
<i>daw-a-toka</i>	PHONETIC EQUIVALENT
<i>w-a-toka</i>	REDUCTION
<i>w-a-rɨk</i>	OTHER REGULAR SOUND CHANGES (Note that [w] = /u/).

4. "THREE" AND THE TRIAL

4.1 THE NUMERAL

The numerals for three are given below:

NTN	<i>kɨsɨl</i>
LEN	<i>kɨsil</i>
WSN	<i>kɨsɨl</i>
NVH	<i>kɨsisɨl</i>
KWM	<i>kahar</i>

With NVH showing reduplication similar to that found in 'two' (and 'five?'), these numerals at first glance present no difficulties, apparently deriving from POC **ra-tolu*.

However, the Tanna languages show one irregularity in this form: the correspondence POC **t* > NTN LEN WSN NVH *s*, KWM *h*. POC **t* > *s* (or *h*) in Tanna usually only before a POC front vowel, whether or not that front vowel is retained:

POC **kati* 'bite' > NTN WSN *us*, LEN *kis*, NVH *as*, KWM *ah/i*;
POC **mate* 'die' > NTN LEN WSN *mis*, KWM *e/mh/a*.

In other environments, POC **t* > NTN WSN *t*, LEN KWM *t* or *r*, and NVH *t* or *l* (the conditions of these last three changes are irrelevant to the present discussion):

POC **tama* 'father' > NTN WSN *tim-*, LEN *rim-*, NVH *lim-*,
KWM *rem-*;

POC **toka* 'dwell' > NTN WSN *a/tiŋ*, LEN *a/rik*, NVH *a/li*,
KWM *a/ra*.

In two cases, however - POC **tolu* 'three' as above, and POC **tokon* 'crutch, pole' > NTN WSN LEN NVH KWM *k-a/skin* - POC **t* unexpectedly behaves as if the following vowel were front, becoming *s* (or *h*) rather than expected *t*, *l*, or *r*. This innovation is shared by all other SNH languages, and I have suggested elsewhere (Lynch 1978) that this may form a significant phonological innovation defining the SNH subgroup.⁹ I have also suggested (Lynch 1976, 1978) an alternative hypothesis that POC may have possessed a palatalized **tʰ*, which had front and back allophones, and that the reflexes *s* and *h* are the POC reflex of PAN **ə* (nowadays written **e*) in the initial position. According to this hypothesis, merger of **tʰ* with **o* (and, in a number of cases, with **e* instead of **ə*) occurred only after the breakup of POC.

Examination of the numerals in other Oceanic languages suggests that this alternative hypothesis has some support. In Lehalurup (Banks Is.) *bi-čael* and Lamenu (Epi) *selu* (< POC **tolu* 'three'), the reflex of **t* is that expected before POC front vowels, and not that expected before back vowels (Tryon 1976). Note also in this connection Apma *ka-čil* and Seke *a-siul* 'three' from Pentecost, for which languages Tryon (1976:12) states that the reflexes *t* and *č* (or *s*) of **t* are unconditioned; and Ranon, Fonah *sul*, Fali *čul*, Baiap *si:*, Sesivi *si*, Port Vato *wari-sie* 'three' from Ambrym, for which languages Tryon (1976:16) says that **t* > *s* (or *č*) "occasionally before POC *o*". These data add more weight to the hypothesis that the first vowel of the POC form for 'three' could well be **ə* rather than **o*.

A further difficulty with the existing reconstruction

concerns the final vowel. I am aware of Blust's (1970) discussion of the frequency of alternations between **i* and **u* in Oceanic. Nevertheless, forms for 'three' with final *i* rather than final *u* are even more common than Blust's discussion would lead us to expect. Although there is no evidence for final *i* from Tanna itself, other SNH languages do show it: Ura on Eromanga has *ge/hli* 'three', while Aneityumese has *e/sej*, where *j* is the regular reflex of **l* before **i*.¹⁰ Note also Hiw (Torres Is.) *bi-toi*, Tambotalo (Santo) *toli*, Port Sandwich (Malekula) *e-rœi*, Dixon Reef (Malekula) *i-tli* (Tryon 1976), and Manam (Papua New Guinea) *toli* (Ross 1977). In addition, all the Central Papuan languages show a form of the type Motu *toi*, although in this case the final *i* appears to be just one example of a regular rule whereby final **olu* > *oi* and final **ulu* > *ui* (Pawley 1975:49ff.).

Pawley's data do allow the suggestion that there may have been some feature of POC **l* which caused a following **u* to become fronted, and for this reason I am at this stage unwilling to suggest that the final vowel be rewritten ambiguously as *(*ui*). Similarly, I do not think enough evidence has as yet been assembled for POC **ə*, and I am thus not prepared at this stage to suggest that **tolu* be rewritten as **təl(ui)*. However, further research may well show that such a revision is necessary.

4.2 THE TRIAL PRONOUN SUFFIX

The Tanna trial pronouns are re-presented here as Table VII.

TABLE VII				
<u>TANNA TRIAL PRONOUNS</u>				
	<i>1st inc.</i>	<i>1st exc.</i>	<i>2nd</i>	<i>3rd</i>
NTN	kitahal	itmahal ¹	itimahal ¹	ilahal
LEN	kathel	kamhel	kamihel ²	ilhel
WSN	kitahal	itimahal ¹	itimwahal ¹	ilahal
NVH	kitasɨl	kimasɨl	kimisɨl	ilisɨl
KWM	kıtahar	kımrahar	kımirahar	irahar

NOTES. 1. See note 2, Table V.
 2. The surface form is *kamhiel* ([kámYɛl]), due to a rule metathesising /h/.

On the basis of these data and comparisons with Table II, the following trial suffixes can be identified:

NTN	-ahal
LEN	-hel
WSN	-ahal
NVH	-asɪl, -sɪl
KWM	-ahar, -rahar

The variation in the NVH suffixes appears to be phonologically conditioned: *-asɪl* after a root ending in a consonant, and *-sɪl* after a vowel-final root. The KWM forms seem a little more complex. I suggest that the alternate *-rahar* may be based on a false analogy: since the dual third person *irau* can be (legitimately) analysed as *i-rau*, it may well be that the trial third person was incorrectly analysed as *i-rahar*, and that *-rahar* was subsequently added to the first exclusive and second person roots.

These trial forms derive from a PTN form **-sVl*, with the initial /a/ in NTN, WSN, KWM, and in one of the NVH allomorphs being best regarded as epenthetic in the environment of a consonant-final root. What is interesting about this form is that the NTN, LEN, and WSN forms at least can *not* derive from a PTN form **-t(ieə)l*, since **t* before a front vowel in these languages becomes *s* and never *h*; the NVH and KWM forms, however, are consistent with either initial **t* (before a front vowel) or initial **s*. Thus although the numeral was inherited with initial **t*, the trial suffix must have been inherited with initial **s*, and we must again reconstruct a suffix which is formally different from the numeral from which, nevertheless, it presumably derives in some way.

While the Aneityum evidence is inconsistent with the PTN trial suffix **-sVl*, it does lend support to the thesis that trial suffixes are different in form from the numeral 'three'. Aneityumese has *e/sej* 'three', suggesting Pre-Aneityumese ***t(ieə)li*, but its trial suffix to pronouns is *-taj*, which derives from Pre-Aneityumese ***-tali*, a form whose first vowel differs from that of the numeral.

Similar evidence for the existence of a trial suffix (often indicating a paucal rather than a true trial) different in phonological form from the numeral 'three' can be found in many Eastern Oceanic (EO) languages. Thus,

Bauan and most PN [=Polynesian] languages exhibit *-tou* for the PEO trial suffix **-tolu*, showing irregular loss of **-l-* ... Wayan has a non-cognate paucal prefix *vati-* (from PEO **pati* 'four'), but *-tou* occurs in all other Western Fijian dialects. In Polynesian languages the type *-tou* marks unlimited plural, the trial/plural contrast having disappeared in PPN. ... The type *-tou*

is found in the Bay of St Philip and St James language, Espiritu Santo, and in at least one group which is probably non-EO: the Nakanai languages of West New Britain. ... [However] the Tongic languages, Tongan and Niuean, show *-tolu* against Nuclear Polynesian *-tou*. (Pawley 1972:122).

Regardless of the phonological differences between the Tanna, Aneityumese, and EO trial suffixes, the process of forming a trial suffix by irregular phonological change in the numeral 'three' emerges clearly. I will discuss this process in more detail in section 6; however, the data presented here strengthen the view proposed first in the previous section, that POC had number affixes which, though probably ultimately derivable from numerals, were phonologically distinct from them.

4.3 THE TRIAL NUMBER-PREFIX TO VERBS

LEN is the only Tanna language which has more than one allomorph of the verbal prefix marking trial subject. The prefixes in the Tanna languages are:

NTN	hl-
LEN	hal-, hai-
WSN	hal-
NVH	lh-
KWM	her-

The LEN trial prefixes have exactly the same distribution as the plural prefixes, and I suggest that this variation developed at a somewhat later stage in LEN on the analogy of the variation in the plural prefix. This is a simpler explanation than assuming that LEN is the only language to retain a PTN distinction which has been lost in all other Tanna languages, since it seems unlikely that four other languages with complex morphophonemics would all level a morphophonemic alternation in just this prefix.

Taking LEN *hal-* as the inherited form and *hai-* as the analogic creation, the PTN trial prefix can be reconstructed as **sVl-*; **V* cannot be **a*, since **s* before **a* becomes *s* and not *h* in NTN, LEN, and WSN. (Metathesis in the NVH form is unexplained). The major point of interest here is that this prefix was probably phonologically identical to the trial pronoun suffix. Recall that the dual prefix and the dual suffix were also phonologically identical. PTN thus appears to have had a single set of number-marking affixes which were used to mark both the number of a pronoun and the number of the subject of a verb.

5. "FOUR" AND THE PLURAL

5.1 THE NUMERAL

The Tanna numerals for 'four' are:

NTN	kuvɨt
LEN	kuvɨr
WSN	kuvɨt
NVH	kuas
KWM	kefa

Now Grace (1969) reconstructed a POC form **pat* 'four', deriving ultimately from PAN **e(m)pat*. This reconstruction derives from Capell, who says (1943:129), however, that "most [South-East Papuan forms] keep the final consonant by supporting it with a developed vowel".

Pawley (1973:173) suggested that this developed vowel - final **i* - is in fact a phonological innovation in POC:

**pati* 'four' shows unanticipated final **-i* in comparison with PAN **e(m)pat* 'four'. **pati* is regularly reflected in many Eastern Oceanic, New Guinea Oceanic and Northwest Melanesian languages, including languages which normally lose PAN final consonants. If it is not cognate with PAN **e(m)pat*, **pati* represents an exclusively shared lexical item.

NVH *kuas* fairly clearly derives from **ka-pati*, for although **p* > NVH *u* is not a regular change, **p* > *u* sporadically in all Tanna languages; moreover, final **t* regularly becomes NVH *s*. Less obviously, KWM *kefa* also derives from **ka-pati*. In KWM, expected final *pVh* becomes *fa*.

KWM *ɛfa*, LEN *ɛpas*, WSN *kuvɨt* 'four';
KWM *nife* 'what?', WSN *paɨ* 'who?'.

The apparent derivation is then **ka-pati* > ***ka-pah* > *kefa*.

The other three languages, however, do not derive from **pati*, but from **pat*. Since **t* before a front vowel regularly becomes *s* in these languages, NTN WSN final *t*, LEN final *r*, cannot derive from a form ending in **ti* (or **te*). I suggest therefore that the northern Tanna forms derive from **ka-pat*, thus justifying Grace's reconstruction.¹¹

PTN, then, appears to have had both **pat* and **pati* as words for 'four'. I suggest that this same situation obtained in POC. Granted that Pawley is correct in reconstructing as POC a form **pati* 'four', there is evidence outside Tanna which supports the reconstruction **pat*. Banks Is. languages, Eromangan, and Mae

are some of the languages which show **t > s* (or *é, h*, etc.) before POC **i*; yet the Banks languages have a form of the type Mota *ni-ḅat*, Eromangan has *d-vat*, and Mae has *i-ḅat*. Similar evidence could be adduced from other New Hebrides languages (Tryon 1976).

In addition, I would like to propose that the Fijian and PN forms for 'four' derive from **pat* and not, as Pawley suggests, from **pati*. Pawley (1972:124) proposed that

PPN **faa* (Walsh and Biggs 1966:6), Bauan, Wayan *vā* 'four' reflect PEO **vati* 'four' [misprint for **pati*: JL] with irregular loss of **-ti* (and compensatory vowel lengthening). ... Gilbertese *a-* 'four' is found, always followed by a classifying suffix. Rotuman *hāke* 'four' has unexplained final *-ke*. Each of the other 26 EO languages [under discussion in Pawley (1972)] regularly reflects the PEO form.

Although Gilbertese has borrowed quite extensively from Polynesian, it is unlikely that Gilbertese *a-* is a borrowing from this source. Loss of final **-ti* has occurred in other nuclear Micronesian languages, ... and is thus probably of considerable antiquity in Gilbertese. ... While no explanation can be offered at present for the Gilbertese form, it is unlikely on other grounds that Gilbertese falls into a subgroup with Fijian and Polynesian.

Pawley thus concludes that this unexplained loss of **-ti* is one innovation marking off the Central Pacific subgroup. The Fijian and Polynesian forms, however, are more easily explainable as deriving from **pat*. Final consonants are regularly lost in these languages, and the forms PPN **faa*, BAU, Wayan *vā*, and the other forms cited are thus derived by regular phonological rule (**-C# > ∅*) without needing to posit such irregular developments as loss of final **-ti*.

The evidence presented above, then, suggests that there were two POC forms for 'four': **pat* and **pati*. Both of these are known to have been in competitive existence in lower-level proto-languages such as Proto-Central Pacific and PTN. I will show in 5.2 and 5.3 below that both of these forms occur even in the same present-day Tanna language (as they do, for example, in Wayan).

5.2 THE PLURAL PRONOUN SUFFIX

The Tanna plural pronouns are re-presented in Table VIII.

TABLE VIII
TANNA PLURAL PRONOUNS

	<i>1st inc.</i>	<i>1st exc.</i>	<i>2nd</i>	<i>3rd</i>
NTN	kitat	itmat ¹	itimat ¹	ilat
LEN	katar	kamar	kamiar	ilar
WSN	kitah	itimah ¹	itimwah ¹	ilah
NVH	kítaua	kímaua	kímia	ilia
KWM	kitaha	kímaha	kímiha	iraha

NOTE. 1. See note 2, Table V.

On the basis of these data and those in Table II we can identify the following plural suffixes:

NTN	-at
LEN	-ar
WSN	-ah
NVH	-aua, -a
KWM	-aha, -ha

Both NVH and KWM have two phonologically conditioned allomorphs: after a consonant, NVH has *-aua* and KWM *-aha*; after a vowel, NVH has *-a* and KWM *-ha*.

A number of Oceanic languages show a plural affix which derives from the POC reconstruction for 'four'. Wayan has already been mentioned above; *wayan* 'four' (< POC **pat*) with the paucal prefix to produce *wayan* 'four' (< POC **pati*) (Pawley 1972:69-73, 122). Capell (1971) also notes this feature in a number of languages: Ithir and Anir actually show a quadral/plural distinction, while other New Ireland languages as well as Tolai (New Britain) and Cao (Ysabel) show a plural pronoun suffix derivable from the POC reconstructions for 'four'.

The NTN and LEN suffixes derive from PTN **-at*, which suggests POC **-pat* with irregular loss of **p*. The WSN suffix derives from **-as*, which not only shows loss of **p* but also unexpected development of the second consonant. The NVH forms *-aua* and *-a* suggest **-pa*, with unexplained loss of **p* after a vowel, and also loss of **t*. The KWM forms are consistent with either **-ati* or **-as* with a supporting final vowel;¹² whatever the origin, the suffix shows loss of **p*.

Thus the Tanna plural suffixes derive from the POC numeral 'four' - NTN and LEN from **pat*, WSN and KWM probably from **pati*, with NVH indeterminate - but, like their dual and trial counter-

parts, they show irregular phonological developments.

5.3 THE PLURAL NUMBER-PREFIX TO VERBS

The prefixes marking plural subject in the Tanna languages are given below:

NTN	ot-
LEN	ar-, ai-
WSN	ot-, oh-
NVH	s-, ha-
KWM	h-

The NTN and KWM forms are invariable, NTN *ot-* deriving from **Vt-* and KWM *h-* from either **ti-* or **s(V)-*.

The WSN forms have the following distribution: *oh-* occurs before any verb beginning with a non-obstruent consonant, while *ot-* occurs before obstruent consonants and vowels. It may be that the base form is *ot-*, with the change to *oh-* being a morphophonemic change developing after the separation of WSN from PTN;¹³ in this case, the WSN prefix would derive from **Vt-*. If this explanation is incorrect, we would have to postulate *ot-* < **Vt-* and *oh-* < **Vs-*.

LEN has *ai-* before verbs beginning with *ua*, *ha*, or *a* when followed by any segment except *i* (with, in the last two cases, concomitant loss of the *a* of the verb root); while *ar-* occurs elsewhere.¹⁴ It is difficult to imagine a natural phonological explanation for deriving *ai-* from *ar-* before *a* and *ua* but not before any other vowel. It is probably more reasonable to assume that *ar-* derives from **Vt-* (ultimately from POC **pat*) and that *ai-* derives from **Vi-* (ultimately from POC **pati* with unanticipated loss of **t*), both cases showing the loss of **p* which occurs in most other Tanna languages in this morpheme.

NVH shows *ha-* before consonants and "potential consonants" and *s-* before non-consonants. Now NVH *h* is the medial reflex of POC medial **-s-*; the form *ha-*, therefore, would derive from **Vs-* with subsequent forward movement of *h* (see note 2, Table VI). Since the initial consonant of the prefix seems to be treated as word-medial (since it is almost always preceded by other prefixes), the NVH prefix *s-* must derive, not from **Vs-*, but from **Vti-*.

Thus the following three prefixes marking plural subject could be reconstructed for PTN:

- 1) **Vt-* > NTN WSN *ot-*, LEN *ar-*.

- 2) *Vs- > NVH *ha-*, WSN *oh-* (if not a subsequent morpho-
 phonemic alternate of *ot-*), KWM *h-* (if not
 < *Vti-).
- 3) *Vti- > NVH *s-* (with loss of *V and *i), LEN *ai-* (with
 loss of *t), KWM *h-* (if not < *Vs-).

Presumably, PTN *Vt- derives from POC *pat; PTN *Vti- from POC *pati; and PTN *Vs- from POC *pati with subsequent reinterpretation of the second consonant. All show unexpected loss of POC *p.

6. CONCLUSIONS

The data presented in this paper show that certain alterations or additions to the POC reconstructed numerals are required. POC *dua 'two' and *lima 'hand, five' remain as reconstructed. Alongside Grace's reconstruction *-kai 'one' we also now have evidence for POC *ta(i) 'one' and POC *nsa 'one'. In section 4 I noted that there is a growing body of evidence to support the reconstruction *təl(ui) as a replacement for POC *tolu 'three', while in section 5 I showed that POC and many of its daughter languages had two competing forms for 'four', *pat and *pati.

I have also demonstrated that the Tanna languages and certain other Oceanic languages show number-marking affixes which seem to be phonologically derivable from the POC numerals, but which differ in form from the numerals in those languages. To recapitulate, Table IX shows the numerals and number-marking affixes that can be reconstructed for PTN. All of the Tanna number-marking affixes differ in phonological form from the corresponding numeral. Many other languages show this same phenomenon in at least one affix; and although the only apparent cognate found so far is PTN *tu-ru* 'dual suffix', it is interesting that, like the Tanna languages, the New Ireland and Tolai languages show occasional loss of the *p of *pat 'four' in the plural (or quadral) suffix.

In discussing the development of these affixes in the Tanna and some other languages, it might be useful to think in terms of a continuum. At one end of the continuum we have languages like Eromangan where, as in English, the plural pronouns refer to *two or more* persons or things. It is only when the number needs to be stressed or emphasised that a numeral is used: thus *koh* 'we inclusive' or *kimi* 'you non-singular' can refer to two, three, or more persons; *koh duru* 'we two inclusive' or *kimi dehel* 'you three', however, are used in much the same way as their English glosses, in cases where it is necessary or desirable to specify the number of people involved.

TABLE IX

PTN NUMERALS AND NUMBER-MARKING AFFIXES

POC		PTN	
		Numerals ¹	Affixes
*ta(i)	'one'	*ta(i)	
*kai	'one'	*ta(i)-kai	
*nsa	'one'	*nsa ²	
*dua	'two'	*dua	*dau
*tolu (*təl(ui)?)	'three'	*t(iə)lV	*sVl ³
*pat	'four'	*pat	*at ⁴
*pati	'four'	*pati	{*ati ⁴ *as ⁴ }
*lima	'five'	*lVm	

- NOTES. 1. The prefix *ka- is omitted.
 2. Only in *ka-nsa-lVm 'five'.
 3. In this word, *V may be any vowel except *a.
 4. The determinate vowel *a can be reconstructed on the basis of the pronoun suffixes, but not on the basis of the verbal prefixes.

Further along the continuum are those languages where the indication of dual and trial seems to be obligatory, but where the form of the dual pronoun is transparently plural pronoun plus numeral. I imagine that most of the EO data given in Pawley (1972) belongs in this category.

At the far end of the continuum belong languages like BAU and those in Tanna. Here the numeral element has to be treated, not as the second lexical morpheme in a compound, but as a suffix to the root. It is at this stage that the suffix begins to undergo morphophonemic changes and loses its phonological resemblance to the numeral, and thus the 'psychological connection' with the numeral is also lost.

It is instructive in this connection to examine changes taking place in the pronoun system of Tok Pisin (New Guinea Pidgin). Early (pidgin) stages of the language showed the following pronoun system:

	<i>singular</i>	<i>plural</i>
1st inc.		yumi
1st exc.	mi	mipela
2nd	yu	yupela
3rd	em	ol

Dual and trial forms were created by inserting the numerals *tu* and *tri* between the pronoun and the suffix: *mitupela* 'we two exclusive', *yutripela* 'you three'. Although other numerals could be used here, in practice these combinations were limited to the numerals 'two' and 'three'.

In the course of time, creole and other fluent speakers of Tok Pisin have contracted the plural suffix to *-pla* (thus *mippla*, *yuppla*); but in addition, they have also contracted the dual form to *mitla* and *yutla*. A truly synchronic analysis would probably be forced to recognise a dual suffix *-t* and a plural suffix *-p*, each of which is followed by a non-singular suffix *-la*. Whatever the analysis, the reinterpretation of the numeral is clear.

The Tok Pisin example vividly illustrates the kinds of changes which I believe have occurred in a number of Oceanic languages. Whether this process was in existence in POC itself is difficult to tell at present, although the fact that BAU and PTN both show a form of **-dau* for the dual suffix, and Tolai, New Ireland languages and PTN show irregular loss of **p* in **pat*/**pati* in the plural suffix suggests that the problem is worth further investigation.

BIBLIOGRAPHY

- Blust, Robert A. 1970. "*i* and *u* in the Austronesian languages". *University of Hawaii Working Papers in Linguistics* 2, 6: 113-145.
- Capell, A. 1943. *The Linguistic Position of South-Eastern Papua*. Sydney.
- 1971. "The Austronesian Languages of Australian New Guinea". In Thomas A. Sebeok (ed.), *Current Trends in Linguistics, vol. 8: Linguistics in Oceania*, pp. 240-340. The Hague.
- Grace, George W. 1969. "A Proto-Oceanic Finder List". *University of Hawaii Working Papers in Linguistics* 1, 1: 39-84.

- Lincoln, Peter C. 1976. *Rai Coast Word List*. Mimeographed, University of Papua New Guinea.
- Lynch, John. 1975. "Lenakel Phonology". *University of Hawaii Working Papers in Linguistics* 7, 1.
- 1976. "A Note on the Proto-Oceanic Vowels". *Kivung* 9, 1: 18-27.
- 1978. *Proto-South Hebridean and Proto-Oceanic*. Paper delivered to the Second International Conference on Austronesian Linguistics.
- Pawley, Andrew. 1972. "On the Internal Relationships of Eastern Oceanic Languages". In R.C. Green and M. Kelly (eds.), *Studies in Oceanic Culture History*, vol. 3 (= Pacific Anthropological Records No. 13), pp. 1-142. Honolulu.
- 1973. "Some Problems in Proto-Oceanic Grammar". *Oceanic Linguistics* 12: 103-188.
- 1975. "The Relationships of the Austronesian Languages of Central Papua: A Preliminary Study". In T.E. Dutton (ed.), *Studies in Languages of Central and South-East Papua*. Pacific Linguistics C29: 3-105.
- 1976. *Some New Oceanic Lexical Comparisons*. Mimeographed, University of Auckland.
- 1977. *On redefining 'Eastern Oceanic'*. Mimeographed, University of Auckland.
- Ross, Malcolm. 1977. *Relationships of the Austronesian Languages of the Sepik and Western Madang Coast of New Guinea*. Mimeographed, University of Papua New Guinea.
- Tryon, D.T. 1976. *New Hebrides Languages: An Internal Classification*. Pacific Linguistics C50.
- Walsh, D.S. and Bruce Biggs. 1966. *Proto-Polynesian Word List I*. Auckland.

NOTES

- ¹ I would like to thank Malcolm Ross and Andrew Taylor for comments on an earlier draft of this paper. I would also like to acknowledge the East-West Center and the University of Papua New Guinea who supported fieldwork in the Southern New Hebrides between 1970 and 1976.

- 2 I also have data on Nivai (Southwest Tanna), which belongs to the same dialect-chain as NVH; the data are not sufficient to be used in this discussion, but show no significant differences from NVH.

LEN has the following phonemes: /pw p t k mw m n ŋ f s l r v w h i e ð a o u/. The other languages have similar phonological systems: WSN is identical except that it may lack /w/; NTN lacks /w/, but has a series of prenasalised stops /mpw mp nt/; NVH lacks /r/ and KWM lacks /l/, but both have /kw/ in place of LEN /w/.

- 3 Pawley (1976:i) warns that "the precise antiquity of many of the reconstructions is uncertain - some must be assigned to Proto-Oceanic ... even on the strictest criteria, while some probably represent post-POC innovations". His evidence for *ka comes from Bauan Fijian, Nggela, Manam, Motu, Roviana, Sa'a, and Tolai.

- 4 This is the PEO as defined by Pawley (1972), and not as subsequently redefined (Pawley 1977).

- 5 Andrew Taylor (pers. comm.) notes that Motu *ta* alone means 'one', and occurs in such expressions as *ruma ta ta* 'each house'; *tamona*, while meaning 'one', contains the morpheme *mo* 'only'.

- 6 Note that initial POC *d- is regularly reflected in NTN LEN WSN as *i*: POC **dan(i)* 'day' > NTN LEN WSN *n/ian*.

- 7 In this case, although the *d in *-*dau* is morpheme-initial, it can never be word-initial and can thus never become *i* in NTN LEN WSN (cf. note 6).

NTN *ao* is the regular reflex of PTN and POC final *-*au*: POC **qau(R)* 'bamboo' > NTN *n/ao*, LEN WSN NVH KWM *nau*.

- 8 Many Tanna languages show epenthesis of an initial vowel (probably POC *a) on to consonant-initial words; for examples, see Lynch (1978).

- 9 Thus POC **tolu* > Aneityumese *de/heī*, Ura *ge/hli* 'three'; POC **tokon* > Aneityumese *de/for* 'crutch'.

- 10 Thus POC **lima* > Aneityumese *ni/jm-* 'hand'; **lipon* > *ni/jho-* 'tooth'; etc. POC **l* does not, however, become *j* when followed by **u*: **luaq* > *a/lou* 'vomit'.

- 11 It is, of course, possible that the NTN, LEN, and WSN forms derive from PTN **ka-pat(aou)*; however, there is no external evidence supporting final *-*a*, *-*o*, or *-*u*.

- 12 KWM occasionally shows a supporting vowel following expected final $-h$: POC **mate* > KWM *e/mh/a* 'die'; **lumut* > *iamh/a* 'seaweed'; and so on.
- 13 Cf. note 1, Table V, where it was shown that *t* before another alveolar in WSN becomes *h*. It may well be that the *h* in *oh-* derives from *ot-* by a rule similar to this.
- 14 The *r* of *ar-* (and also the *l* of *hal-*) is deleted when the following consonant is alveolar; cf. note 1, Table V.