

18 *At sixes and sevens: the development of numeral systems in Vanuatu and New Caledonia*

JOHN LYNCH

1 Introduction¹

In a review of Lynch, Ross and Crowley's (2002) *The Oceanic languages*, Robert Blust went on a self-confessed 'major digression' (2005:556). He directed attention to 'Vanuatu and southern Melanesia [New Caledonia and the Loyalties], where Papuan languages are absent, but Papuan features such as [quinary numeral systems and serial verb constructions] are present' (554), and concluded that 'the Papuan features of language, culture, and physiognomy that are common to [Austronesian] speakers in Vanuatu and southern Melanesia must have been acquired by contact *in situ*' (555)—in other words, by contact with Papuan languages which must once have been spoken in that area but which have since presumably died out.

In this paper, I investigate the nature of the numeral systems of the languages of Vanuatu and New Caledonia. I should point out from the outset that this paper is *not* an attempt to prove Blust right or wrong. It may well be that the structural shift from the purely decimal system of Proto Oceanic to systems based on 5 in many of these languages was due to Papuan influence. However, there is not a single shred of evidence to suggest that this was the case, since there are no Papuan languages spoken in this area, and no evidence—either that I am aware of or advanced by Blust—that there ever were. Rather, what this paper does is to attempt to outline the morphological processes involved in the development of those systems which are not purely decimal. Given the explanations I propose, it *may* be possible to find extant Papuan languages in the Pacific which could form the basis for calques which arose in Vanuatu and New Caledonia.

¹ It is a pleasure to be able to offer this paper in honour of Malcolm Ross, whose contribution to Austronesian and Papuan linguistics over the last two decades have been enormous. Malcolm made a somewhat late entry into these fields, stimulated by the diverse linguistic environment during his time as an English master at Keravat National High School and later Principal of Goroka Teacher's College, both in Papua New Guinea, during the 1980s. His mid-career shift in interest, however, has benefited Austronesian and Papuan linguistic research tremendously.

I thank Isabelle Bril, Michael Dunn, Robert Early, Peter Lincoln, Claire Moyse, Åshild Næss and Hannah Vari-Bogiri for supplying data and/or interpretations of data from various languages; and Beth Evans, Alexandre François and Peter Lincoln for comments on an earlier draft. Responsibility for errors and misinterpretations is mine alone.

Following earlier studies (e.g. Tryon 1976, Clark 1985), I use the following terms to describe different kinds of systems:²

- DECIMAL refers to systems in which the numerals 1–10 are monomorphemic (or have monomorphemic roots with a synchronic or fossilised prefix), and where 20 is represented by a compound involving 2 and 10.
- IMPERFECT DECIMAL refers to systems which differ from decimal systems only in that the numerals 6–9 are compounds generally involving the numerals 1–4 in some way. In these systems, 10 and 20 are constructed in the same way as for decimal systems.
- QUINARY refers to systems in which there are no monomorphemic numerals above 5; 10 is either 5-(and)-5 or ‘two fives/hands’; and 20 is a compound of the type ‘one person’.³
- MIXED refers to systems with features of both imperfect decimal and quinary systems. Typically, 10 is ‘two fives/hands’, but 20 is ‘two tens’.

This paper deals mainly with the numerals 1–10, though I will briefly mention forms for 20 and make mention in passing of higher numerals.

The ancestor of the languages of this area, Proto Oceanic, had a decimal system (see, for example, (1) below). Where imperfect decimal, quinary and mixed systems have developed, they involve the replacement of monomorphemic forms with analytical forms. I will therefore use the term INNOVATIVE to refer to imperfect decimal, quinary and mixed systems when they do not need to be distinguished from each other, and the term LIGATURE (abbreviated LIG) to refer to elements which conjoin numerals in compounds (e.g. Tolomako *lina-rave-rua* 5-LIG-2 ‘seven’). Arabic numerals as glosses appear without quotation marks.

2 Distribution of decimal and innovative systems

Full decimal systems are found in the following areas in Vanuatu–New Caledonia:

- all the languages of Ambae;
- Raga in the north of Pentecost;
- the Cape Cumberland area of the extreme northwest of Santo and, at the opposite end of the island, the offshore islands of Malo and its neighbours; and
- the northern part of Malakula, adjacent to the Malo area.

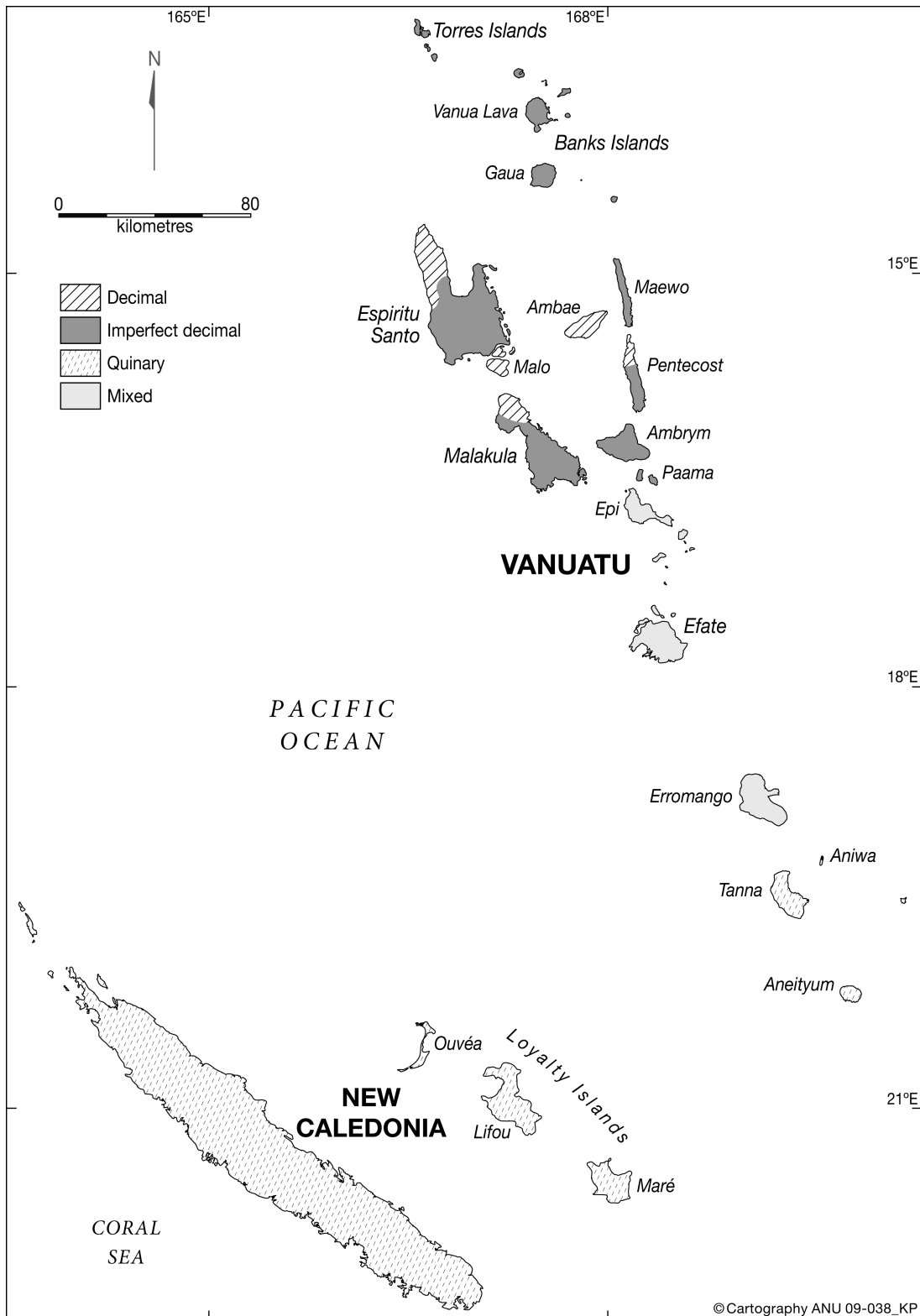
A sample is given in (1).⁴

Conversely, innovative systems are found in much of Northern Vanuatu, almost all of Central Vanuatu, and throughout the two more southerly groups. Map 1 illustrates the geographical location of each type of system.

² These terms, or terms like them, were used by Tryon (1976) and Clark (1985) in their studies of Vanuatu languages. They differ slightly from terms used by, for example, Comrie (2008); but these differences are not significant.

³ Lincoln (in press) notes that so-called quinary systems are not really quinary or base-5 in the mathematical sense, since the ‘milestones’ are not 5, $5^2 = 25$, $5^3 = 125$, etc. However, this term has been so widely used in Oceanic studies that I will retain it here.

⁴ Sources of language data and comments on the orthography used can be found in the appendix. POc - Proto Oceanic, PEOc - Proto Eastern Oceanic; PNCV - Proto North-Central Vanuatu.



Map 1: Geographical location of numeral systems in Vanuatu and New Caledonia

(1)	POc	Raga	Araki	Uripiv
		Pentecost	Santo	Malakula
1	various ⁵	<i>tea</i>	<i>hese</i>	<i>san(san)</i>
2	*rua	<i>rua</i>	<i>dua</i>	<i>e/ru</i>
3	*tolu	<i>tolu</i>	<i>rolu</i>	<i>i/tul</i>
4	*pat(i)	<i>vasi</i>	<i>v'ari</i>	<i>i/vij</i>
5	*lima	<i>lima</i>	<i>lim'a</i>	<i>e/lim</i>
6	*onom	<i>ono</i>	<i>hai/ono</i>	<i>ou/won</i>
7	*pitu	<i>bitu</i>	<i>hai/p'iru</i>	<i>e/but</i>
8	*walu	<i>v^welu</i>	<i>ha/ualu</i>	<i>o/wil</i>
9	*siwa	<i>siva</i>	<i>hai/sua</i>	<i>e/siw</i>
10	*sa-[ŋa]-puluq	<i>haŋvul(u)</i>	<i>saŋavulu</i>	<i>esŋavəl</i>
20	*rua-ŋa-puluq	<i>ŋavul-gai-rua</i>	<i>ŋavul dua</i>	<i>ŋavəl eru</i>

2.1 Innovative systems in Vanuatu and mainland New Caledonia

Innovative systems in Vanuatu and New Caledonia (excluding the Loyalties, discussed in the next section) can be subdivided into a number of different types, using different features to subclassify them. A sample of these systems in northern and central Vanuatu is given in (2) and in southern Vanuatu and New Caledonia in (3). As discussed in §1, innovative forms for the numerals are often compounds which include a form which is the same as, or is related to, the numerals 1–4. In (2) and (3) below, bolding marks the numerals 1–4 in the forms for 6–9.⁶ Of the languages given here, Merei and Naman are imperfect decimal, Lewo and Sye are mixed, and the remainder are quinary.

(2)	POc	Merei	Naman	SE Ambrym	Lewo
		Santo	Malakula	Ambrym	Epi
1	various	<i>ese</i>	<i>savakh</i>	<i>tei</i>	<i>taŋa</i>
2	*rua	<i>ruwa</i>	<i>iru</i>	<i>lu</i>	<i>lua</i>
3	*tolu	<i>tolu</i>	<i>itəl</i>	<i>tol</i>	<i>telu</i>
4	*pat(i)	<i>vat</i>	<i>i</i>	<i>ves</i>	<i>hat vari</i>
5	*lima	<i>lima</i>	<i>iləm</i>	<i>lim</i>	<i>lima</i>
6	*onom	<i>maravo</i>	<i>nsous</i>	<i>tehesav</i>	<i>orai</i>
7	*pitu	<i>ravorua</i>	<i>nsuru</i>	<i>luhesa</i>	<i>olua</i>
8	*walu	<i>raptol</i>	<i>nsutəl</i>	<i>olhesa</i>	<i>orelu</i>
9	*siwa	<i>raitat</i>	<i>nsoves</i>	<i>hathesa</i>	<i>ovari</i>
10	*sa-[ŋa]-puluq	<i>saŋavul</i>	<i>saŋavəl</i>	<i>he-xa-lu</i>	<i>lua-lima</i>
20	*rua-ŋa-puluq	<i>ŋavul-rua</i>	<i>(na)ŋavəl iru</i>	<i>hanu tap</i>	<i>lua-lima yam lua</i>

⁵ Various forms have been reconstructed with the meaning 1 in POc, including *ta, *sa, *tai, *kai, and various combinations of these. In many languages, most numerals (though often not 1 and 10) have a historical or synchronic prefix, which often is, or is derived from, a 3SG or 3NSG verbal prefix: Uripiv in (1) illustrates this.

⁶ Some relationships may not be immediately obvious: Lewo, for example, reflects POc *t as *t* initially (cf. 1, which is probably a compound, and 3) but as *r* medially (as in 6 and 8); the Southeast Ambrym form for 8 unexpectedly loses initial *t*; and so on.

The following comments can be made in relation to the data presented in (2) and (3). The form of the numerals 6–9 (or 7–9, since 6 is sometimes aberrant—see §3) may be:

- LIG-numeral (Merei, Naman, Lewo);
- numeral-LIG (Southeast Ambrym); or
- 5-(LIG)-numeral (Sye, Lenakel, Nemi, Xârâcùù).

The form of the numeral 10 may be:

- a synchronic monomorphemic form, generally a reflex of *sa-[ŋa]-puluq (Merei, Naman);
- a historical multiplicative, involving terms for 2 and 5 or ‘hand’ (Southeast Ambrym,⁷ Lewo, Sye, Xârâcùù?);
- 5-(LIG)-5 (Lenakel); or
- an apparent compound, whose etymology is however unclear (Nemi?).

The form of the numeral 20 may be:

- a multiplicative involving terms for 2 and 10 (Merei, Naman, Lewo, Sye); or
- a phrase incorporating the word for ‘man’, ‘person’, or ‘fingers/toes/digits’ (Southeast Ambrym, Lenakel, Nemi, Xârâcùù). Lenakel *ieramim karena raka*, for example, is literally man one he-is-not; Southeast Ambrym *hanu tap* is person whole; etc.

(3)	POc	Sye	Lenakel	Nemi	Xârâcùù
		Erromango	Tanna	N. New Cal.	S. New Cal.
1	various	<i>hai(teven)</i>	<i>karena</i>	<i>heec</i>	<i>chaa</i>
2	*rua	<i>nduru</i>	<i>kiu</i>	<i>heluk</i>	<i>baaru</i>
3	*tolu	<i>ndehel</i>	<i>kasil</i>	<i>heyen</i>	<i>bachéé</i>
4	*pat(i)	<i>ndvat</i>	<i>kuvər</i>	<i>phoec</i>	<i>kêrêfùè</i>
5	*lima	<i>sukrim</i>	<i>katilum</i>	<i>nim</i>	<i>kêrênürü</i>
6	*onom	<i>mehikai</i>	<i>katilum-karena</i>	<i>ni-bweec</i>	<i>kêrênürü mē chaa</i>
7	*pitu	<i>sukrimnduru</i>	<i>katilum-kiu</i>	<i>ni-bweluk</i>	<i>kêrênürü mē baaru</i>
8	*walu	<i>sukrimndehel</i>	<i>katilum-kasil</i>	<i>ni-bweyen</i>	<i>kêrênürü mē bachéé</i>
9	*siwa	<i>sukrimndvat</i>	<i>katilum-kuvər</i>	<i>ni-bovac</i>	<i>kêrênürü mē kêrêfùè</i>
10	*sa-[ŋa]-puluq	<i>narwolem</i>	<i>katilum-katilum</i>	<i>paidu</i>	<i>duchêêxê</i>
20	*rua-ŋa-puluq	<i>narwolem</i> <i>nduru</i>	<i>Ieramim karena</i> <i>raka</i>	<i>hee kahok</i>	<i>xê chaa kamûrû</i>

2.2 Loyalty Islands

The Loyalty Islands languages follow a somewhat different pattern. The numerals 1–20 in the three Loyalties languages are given in (4).

These three languages show some similarities and some differences:⁸

⁷ For Southeast Ambrym, Parker (1970) gives *hexalu* 10 (cf. *he-* ‘hand’, *lu* 2) and *hanu tap* (person whole) 20. Some modern speakers however have reanalysed this system, and give *hanutap tei* 10 and *hanutap lu* 20. The *he* in the ligature *-hesa(v)* in 6–9 is probably ‘hand’.

⁸ I am grateful to Claire Moysé for providing the Nengone data and assisting with much of the analysis of Drehu and Nengone.

- **Iaai** has two forms for 5 which appear to be in free variation. It has something which could roughly be translated as ‘two hands’, ‘three hands’ and ‘four hands’ for 10, 15 and 20 (*baa-* being a form which is used for naming fingers, though not the actual form for ‘hand’, which is *beñi-*). For 10 there is the alternative compound *li beñi-ta* (DUAL hand-1PL.INCL). Other numerals show the numerals 1–4 linked to *thabüŋ* 5, *li beñi-ta* 10 and *baa-kun* 15 by *ke nua* ‘and again’.
- In **Drehu**, the numerals 5, 10 and 15 are formed on the base *-pi*, which may possibly be related to a form *pi* meaning ‘last, after’. Although the meaning of *trii* in *trii-pi* 5 is unknown, 10 and 15 are *2-pi* and *3-pi*. Drehu is particularly interesting, not only in prefixing the numeral to the ligature, but also in having three separate ligatures: *-ŋömen* ~ *-ŋemen* for 6–9, *-ko* for 11–14, and *-qaihana* for 16–19. Parts of these ligatures can be identified: the *ŋe* of *-ŋömen* ~ *-ŋemen* is ‘and’, *ko* is ‘towards’, and the *qai* of *-qaihana* means ‘coming from’. 20 is ‘one person’.
- **Nengone** has 2×5 or ‘two hands’ for 10; 11–14 are additives to both of these (*ne* = ‘and’). 15 = ‘two hands and five again’, 16–19 are additives to this (minus *yawe* ‘again’), and 20 is ‘one person’.

(4)	Iaai	Drehu	Nengone	
1	<i>xaca</i>	<i>caa(s)</i>	<i>sa</i>	
2	<i>lo</i>	<i>lue</i>	<i>rewe</i>	
3	<i>kun</i>	<i>köni</i>	<i>ten(e)</i>	
4	<i>væk</i>	<i>eke</i>	<i>ece</i>	
5	<i>baa-xaca, thabüŋ</i>	<i>trii-pi</i>	<i>sedoŋ</i>	
6	<i>thabüŋ ke nua xaca</i>	<i>caa-ŋömen</i>	<i>sedoŋ ne sa</i>	
7	<i>thabüŋ ke nua lo</i>	<i>lue-ŋömen</i>	<i>sedoŋ ne rewe</i>	
8	<i>thabüŋ ke nua kun</i>	<i>köni-ŋömen</i>	<i>sedoŋ ne ten</i>	
9	<i>thabüŋ ke nua væk</i>	<i>eke-ŋömen</i>	<i>sedoŋ ne ece</i>	
10	<i>baa-lo, li beñi-ta</i>	<i>lue-pi</i>	<i>rue sedoŋ</i>	<i>rue tubenin</i>
11	<i>li beñi-ta ke nua xaca</i>	<i>caa-ko</i>	<i>rue sedoŋ ne sa</i>	<i>rue tubenin ne sa</i>
12	<i>li beñi-ta ke nua lo</i>	<i>lue-ko</i>	<i>rue sedoŋ ne rewe</i>	<i>rue tubenin ne rewe</i>
13	<i>li beñi-ta ke nua kun</i>	<i>köni-ko</i>	<i>rue sedoŋ ne ten</i>	<i>rue tubenin ne ten</i>
14	<i>li beñi-ta ke nua væk</i>	<i>eke-ko</i>	<i>rue sedoŋ ne ece</i>	<i>rue tubenin ne ece</i>
15	<i>baa-kun</i>		<i>köni-pi</i>	<i>rue tubenin ne sedoŋ yawe</i>
16	<i>baa-kun ke nua xaca</i>	<i>caa-qaihana</i>	<i>rue tubenin ne sedoŋ ne sa</i>	
17	<i>baa-kun ke nua lo</i>	<i>lue-qaihana</i>	<i>rue tubenin ne sedoŋ ne rewe</i>	
18	<i>baa-kun ke nua kun</i>	<i>köni-qaihana</i>	<i>rue tubenin ne sedoŋ ne ten</i>	
19	<i>baa-kun ke nua væk</i>	<i>eke-qaihana</i>	<i>rue tubenin ne sedoŋ ne ece</i>	
20	<i>baa-væk</i>	<i>caatr</i>	<i>sa reŋom</i>	

2.3 Distribution of different types of innovative systems

The distribution of these various subtypes is given in Table 1.⁹ Leaving aside some minor details for the moment, it seems that we can divide these languages into two major

⁹ Single-language abbreviations in Table 1 are: for SE Ambrym/Paama, P - Paamese and SEA - Southeast Ambrym; for Erromango, S - Sye, Ur - Ura and Ut - Utaha; and for the Loyalties, D - Drehu, I - Iaai and N - Nengone. Otherwise, ✓ indicates that this feature is found in all languages in the island(s) concerned.

groups on the basis of the nature of the innovative system: a northern area comprising all of the Northern Vanuatu and some of the Central Vanuatu subgroups, where the systems are imperfect decimal, and a southern area comprising most of Southern Vanuatu and all of New Caledonian, where the systems are quinary. There is a third, intermediate group separated from the northern and southern groups by solid lines in the table, in which mixed systems occur. Shading in the northern and southern groups indicates features typical of each group. The symbol † beside the name of an island or island group indicates that conservative decimal systems are also found there.

Table 1: Distribution of different innovative systems

	NUMERALS 6–9			FORM FOR 10			FORM FOR 20	
	5-(LIG)- num	LIG-num	num- LIG	single morph	5-(LIG) -5	2 × 5/ hand	2 × 10	person, digits
IMPERFECT DECIMAL								
Torres/Banks		✓		✓			✓?	
Maewo		✓		✓			✓?	
Pentecost †		✓		✓			✓?	
Santo †	few	most		✓			✓?	
Malakula †		✓		✓			✓?	
N and W Ambrym		✓		✓			✓?	
MIXED								
Epi-Efate		✓				✓	✓?	
Erromango	S	Ur, Ut				✓	✓	
QUINARY								
SE Ambrym/ Paama		P	SEA			✓		✓
Tanna	✓				✓			✓
Aneityum		✓			✓			✓?
N. New Cal.	✓					✓?		✓
S. New Cal.	✓					✓?		✓
Loyalties	I, N		D			✓		D, N; I = 4×hand

The features of the two major areas are as follows. In the northern, imperfect decimal, area (which also houses some languages with decimal systems):

- forms for 6–9 are almost universally LIG-numeral;
- the form for 10 is a single morpheme (synchronically), and almost everywhere derives from POC *sa-[ŋa]-puluq; and
- the form for 20 is a multiple of 2 and 10.

In the southern, quinary, area:

- forms for 6–9 are almost universally 5-LIG-numeral;
- the form for 10 is normally either 5-LIG-5, 2×5 or $2 \times$ ‘hand’; and
- the form for 20 is normally ‘one person’, ‘person’s digits’, or some other phrase referring to a person, implying that the counting of ten fingers and ten toes is complete.

The intermediate area in the middle of Table 1, incorporating Epi, the Shepherds, Efate and Erromango, shows a mixed system and seems to be a transitional region between the northern and southern areas; it has no single morpheme for 10, but does show ‘two tens/hands’ for 20.

2.4 Incipient innovative systems?

There is a group of contiguous languages—Araki in southern Santo and five or six languages in northern Malakula—which are clearly decimal, with reflexes of *onom 6, *pitu 7, *walu 8 and *siwa 9. In these languages, forms for 1–5 are either historically monomorphemic or else have a prefix which is or was the 3SG subject marker. However, the numerals 6–9 take a quite different prefix. I present data below from Araki and from five northern Malakula languages.

(5)	POc	Araki	Nese	Botovro	Tirax	Malua Bay	Vovo
1	various	<i>hese</i>	<i>sakhal</i>	<i>hual</i>	<i>hkhal</i>	<i>səkhal</i>	<i>hual</i>
2	*rua	<i>dua</i>	<i>rru</i>	<i>rue</i>	<i>ru</i>	<i>i/ru</i>	<i>rue</i>
3	*tolu	<i>rolu</i>	<i>til</i>	<i>til</i>	<i>til</i>	<i>i/til</i>	<i>til</i>
4	*pat(i)	<i>v’ari</i>	<i>v’at</i>	<i>v’at</i>	<i>vat</i>	<i>i/vat</i>	<i>vat</i>
5	*lima	<i>lim’a</i>	<i>line</i>	<i>lim’e</i>	<i>lin</i>	<i>i/ləm</i>	<i>lime</i>
6	*onom	<i>hai/ono</i>	<i>kh/on</i>	<i>h/on</i>	<i>khɔ/wɛn</i>	<i>kho/en</i>	<i>on</i>
7	*pitu	<i>hai/p’iru</i>	<i>kho/dit</i>	<i>huo/dit</i>	<i>khə/dit</i>	<i>i/bit</i>	<i>kho/dit</i>
8	*walu	<i>ha/ualu</i>	<i>kho/al</i>	<i>ho/al</i>	<i>khɛ/wɛl</i>	<i>o/wel</i>	<i>kho/al</i>
9	*siwa	<i>hai/sua</i>	<i>khe/sve</i>	<i>khe/sve</i>	<i>khɛ/siv</i>	<i>kha/səp</i>	<i>khe/hive</i>
10	*sa-[ŋa]-puluq	<i>saŋavulu</i>	<i>saŋav’il</i>	<i>haŋavul</i>	<i>hŋavil</i>	<i>saŋavil</i>	<i>haŋavil</i>

The forms for 6–9 in these languages contain a prefix which would derive from *kV-, whose possible origin will be discussed in more detail in §4.2.¹⁰ In these languages, then, 6–9 are morphologically marked, in comparison with the unmarked 1–5. They thus resemble languages with imperfect decimal systems in treating numerals above 5 differently from those below it, even though the POc roots for 6–9 are retained. Whether this represents some kind of influence from neighbouring imperfect decimal languages, or the beginning of a change from decimal to imperfect decimal, is difficult to tell.

2.5 Higher numerals

Higher numerals—by which I mean monomorphemic forms for 100, 1000 and the like—are found in a number of Oceanic languages with decimal systems. However, a number of languages in the northern part of Vanuatu with imperfect decimal systems also have forms for 100 and 1000 which are not compounds or paraphrases (though in the

¹⁰ Araki *h* is the regular reflex of *k (e.g. *huru* ‘louse’ < *kutu, *siho* ‘kingfisher’ < *siko).

southern part of Vanuatu and in New Caledonia 100, for example, is usually ‘five men’). For example:

(6)		100	1000		100	1000
	Mwotlap	<i>m̃eldēl</i>	<i>tey</i>	Uripiv	<i>nuʒut</i>	<i>evin, nuvesan</i>
	Mota	<i>m̃elnol</i>	<i>tar</i>	Neverver	<i>naʒat</i>	<i>netar</i>
	Merei	<i>lavul</i>		Namakir	<i>na-ponati</i>	
	Avava	<i>aʒat</i>	<i>atar</i>	Nguna	<i>ṽonotia</i>	<i>maanu</i>
	Neve’ei	<i>naʒat</i>	<i>netar</i>	S Efate	<i>tifli</i>	<i>p̃ on</i>
	Naman	<i>noʒot</i>		Sye	<i>nalem</i>	
	Tape		<i>itar</i>			

The Mwotlap and Mota forms for 100 are historically bimorphemic: François (2005:498) reconstructs an earlier *m̃ele-dolu 100 (lit. cycas palm-whole), with the second element probably deriving from PEOc *udolu ‘very many, all’ (Pawley 1972:55). The Malakula *ʒat-type forms for 100 may derive from Proto Malayo-Polynesian *sa-ʒa-Ratus 100 (Robert Blust, pers. comm. 4 Feb. 2009) with regular loss of *sa- and the final syllable (and fairly regular loss of *R): thus *sa-ʒa-Ratus > -ʒat. For 1000, PEOc *tari, which occurs in other languages meaning ‘(very) many’, can be reconstructed (Pawley 1972:55; see also François 2005:500).

3 Forms for ‘six’

Forms for 6 in innovative systems do not always follow the same pattern as forms for 7–9, and it will be worth discussing these briefly here, both in view of the discussion in §2 and also as an introduction to the discussion on ligatures in §4. Two examples of ‘irregular’ forms for 6 can be found in (2) and (3) above: Merei has *ma-ravo* (abbreviated form of 5 + LIG) for 6 but (underlying) *ravo* (LIG) + 2/3/4 for 7/8/9; and Sye has *mehikai* 6, but 7/8/9 are compounds of *sukrim* 5 + 2/3/4.¹¹

A sample of ‘irregular’ forms for 6 is given in (7), and these are discussed in more detail below.

(7)	Wusi	Tolomako	Tasiriki	Tape	Avava	Ura	Anejom̃
	Santo	Santo	Santo	Malakula	Malakula	Erromango	Aneityum
1	<i>ehe</i>	<i>tea</i>	<i>ʔese</i>	<i>isig</i>	<i>sapm</i>	<i>sai</i>	<i>ithii</i>
5	<i>lima</i>	<i>lina</i>	<i>lima</i>	<i>iləm</i>	<i>ilim</i>	<i>suworem</i>	<i>mijman</i>
6	<i>lima-rave</i>	<i>lina-rave</i>	<i>ha-ʔese</i>	<i>ləm-jis</i>	<i>sout</i>	<i>misai</i>	<i>meled</i>
7	<i>rave-rua</i>	<i>lina-rave-rua</i>	<i>ravaʔa-rua</i>	<i>ji-ru</i>	<i>sou-ru</i>	<i>sinelu</i>	<i>meled-erou</i>
8	<i>rap-tol</i>	<i>lina-rave-tolu</i>	<i>ravaʔa-tolu</i>	<i>ji-təl</i>	<i>se-tl</i>	<i>sinehli</i>	<i>meled-esej</i>
9	<i>ra-pati</i>	<i>lina-ra-tati</i>	<i>ravaʔa-vati</i>	<i>ji-vet</i>	<i>sa-vat</i>	<i>sinivat</i>	<i>meled-emanohowan</i>

¹¹ I exclude from the ‘irregular’ category forms for 6 which basically follow the same pattern as 7–9 (in being composed of 5-1, 5-LIG-1 or LIG-1) but where the element representing 1 is a recognisably reduced or modified form of the numeral 1, as in Southeast Ambrym and Lewo in (2), or some other form meaning ‘one’ which is not used in serial counting and one or more others used within noun phrases.

Three different types of irregular forms for 6 are found in Santo, illustrated in (7) by Wusi, Tolomako and Tasiriki.

- The Wusi-type has 5-LIG for 6 but LIG-2/3/4 for 7–9. This type is found in most languages of Santo which have imperfect decimal systems, including various dialects of Central Santo and Southwest Santo, Kiai, Shark Bay, Mav’ea and Merei—i.e. much of the southern half of the island as well as the northeast.
- The Tolomako-type also has 5-LIG for 6, but 5-LIG-2/3/4 for 7–9. This type is also found in Sakao, which neighbours Tolomako in the Big Bay area, but in addition it occurs further to the south in South Central Santo.
- In the Tasiriki-type, 6–9 are all of the form LIG-1/2/3/4, but the ligatures are phonologically unrelated: Tasiriki *haʔ*-, Akei *awe*- with 6, but Tasiriki *ravaʔa*-, Akei *rava*- with 7–9. (Tasiriki and Akei are both dialects of Southwest Santo, though other dialects are of the Wusi-type).

Two different types of ‘irregular’ forms for 6 are found in Malakula, illustrated in (7) by Tape and Avava.

- Tape and V’ënen Taut (Big Nambas) have 5-1 for 6 but LIG-2/3/4 for 7–9.
- The Avava-type has the same LIG morpheme (given some vowel alternations) for all numerals 6–9. However, the form following LIG in 6 bears no phonological resemblance to the form for 1, nor to any other form with a similar meaning: Crowley (2006a:57) says of Avava *sout* 6 that it ‘involves the same initial element *sV*- noted for the numerals 7–9, but with a following element that cannot be related to any of the other cardinal numerals’. Note Naman *nsous* in (2), Avava *sout* in (7) and the following in two other Malakula languages:

(8)	Neve’ei	Neverver		Neve’ei	Neverver
	1	<i>sevakh skham</i>		6	<i>nsouh jos</i>
	2	<i>iru ru</i>		7	<i>nsu-ru jo-ru</i>
	3	<i>itl tl</i>		8	<i>nsu-tl jo-tl</i>
	4	<i>ivah vas</i>		9	<i>nsa-vah jo-vas</i>
	5	<i>ilim lim</i>		10	<i>naṇavil naṇavul</i>

In Paamese, the form for 6 uses a different LIG from the one used with 7–9: note *lahi-tāi* LIG-1 ‘six’, but *lau-lu*, *lau-tel* and *lau-hat* 7–9. Crowley (1982:98) suggests that the ligature for 6 is the verb *lahi* ‘carry’ and the other ligature is the noun *lau*- ‘leaf’. I will return to this in the next section.

In Erromango, forms for 6 seem to derive from *ma ‘and’ + 1 (although Sye *mehikai* does not bear too close a resemblance to *hai(teven)* 1). Forms for 7–9 are based on *sukrim* 5 + 2–4 in Sye, but on LIG-2–4 in the other Erromangan language, as illustrated in (9).¹²

¹² Ura and Utaha have lost *pat(i) for the numeral 4 and replaced it with an additive form 2-and-2. However, a reflex of *pat(i) shows up in the compound form 9.

(9)	Ura	Utaha	Ura	Utaha
	1 <i>sai</i>	<i>so voi</i>	6 <i>misai</i>	<i>mise vai</i>
	2 <i>gelu</i>	<i>kalu</i>	7 <i>sine-lu</i>	<i>simna-lu</i>
	3 <i>gehli</i>	<i>kihili</i>	8 <i>sine-hli</i>	<i>simni-heli</i>
	4 <i>lemelu</i>	<i>lemelu</i>	9 <i>sini-vat</i>	<i>simni-vat</i>
	5 <i>suworem</i>	<i>sukrim</i>	10 <i>lurem, durem</i>	<i>narolem</i>

The Anejoñ form for 6 is just LIG: *meled* 6 is composed of *m-* ‘echo subject prefix’ + *eled* ‘be left over’. Numerals 7–9 are composed of *meled*-2/3/4.¹³

Forms for 6 thus often pattern differently from forms for 7–9. This is an important element in the hypothesis of the development of the numeral systems which I will develop in §4 and §5.

4 Ligatures

The ligature used in those quinary languages in the southern area which do not simply conjoin 5 and another numeral is either a reflex of POC **ma* ‘and’, some other coordinating conjunction, or a phrase like ‘and more’ or ‘and again’. This is not of any particular interest. Of more interest historically is the type of ligature used in the northern languages, where the numeral 5 normally does not appear (except in 6 and in the Tolomako-type described above), and where the ligature is *not* a coordinating conjunction. Three reasonably widespread ligatures can be identified.

4.1 PNCV **lave-a*

The most widespread ligature in Vanuatu has the form *lavV-* or *ravV-*. Cognates can be seen in Merei in (2) and in Wusi, Tolomako and Tasiriki in (7). The distribution of related forms is as follows:

- all Torres, Banks and Maewo languages;
- all Pentecost languages (except for the decimal Raga);
- all Santo languages which have imperfect decimal systems;
- some Malakula languages which have imperfect decimal systems: Axamb, Lendamboi, Maskelynes, Banam Bay, Unua and Aveteian. These are spoken mainly in south and east Malakula, though there is no particular genetic connection between just these languages, as they belong to two different first-order subgroups of Malakula languages (Lynch 2007), and some of their close relatives behave differently; and
- Paamese.

In addition, the following are probably also related:

- the type *milip-* or *miliv-* in a number of Ambrym languages, which may represent the *lavV-* form preceded by a subject-TAM marker, or by a conjunction, or possibly by a reduced form of **lima* 5; and
- the form *la-* in the languages of the Shepherds and Efate, in which the second syllable was lost.

¹³ Anejoñ *d* is a voiceless interdental fricative /θ/, and *eled* appears to be related to PNCV **malazi* ‘leftover food’ (with probably a stative prefix **ma-*).

Some additional examples of the forms for 6–9 in one language from each of these areas follow:

(10)	Mwotlap	Peterara	Sowa	Kiai	Lendamboi	Fali	Nguna
	Banks	Maewo	Pentecost	Santo	Malakula	Ambrym	Shepherds
6	<i>leve-te</i>	<i>lav-tu^vale</i>	<i>lo-wal</i>	<i>lima-rave</i>	<i>i-lav-sua</i>	<i>milip-t^vfe</i>	<i>la-tesa</i>
7	<i>livi-yō</i>	<i>lav-rua</i>	<i>lew-ru</i>	<i>rav-ua</i>	<i>i-lav-rua</i>	<i>milip-ru</i>	<i>la-rua</i>
8	<i>leve-tēl</i>	<i>lav-tolu</i>	<i>lep-tul</i>	<i>ra-tolu</i>	<i>i-lavu-təl</i>	<i>milip-t^vfil</i>	<i>la-tolu</i>
9	<i>leve-vet</i>	<i>lei-vati</i>	<i>lak-pat</i>	<i>ra-pati</i>	<i>i-la-vas</i>	<i>milip-fer</i>	<i>lo-veti</i>

There is a certain amount of erosion of the second syllable in a number of areas, and also a certain amount of variation in the vowels, much of it possibly due to some kind of assimilation to the vowel of the numeral root. Crowley's comment in §3 above regarding Paamese is a case in point: though he suggests that *lahi-* used with 6 and *lau-* with 7–9 have different meanings/origins, I suggest that they both derive from the same form, with loss of *v and a vowel change in the latter. The examples in (10) also illustrate these kinds of changes: Mwotlap shows vowel alternations, Peterara has loss of *v in 9, Sowa has vowel and consonant changes, Lendamboi loses *v in 9, and Nguna shows loss of the second syllable plus vowel changes.

The initial consonant reflects *l in most of the languages in which this occurs, although some Santo languages suggest initial *r rather than *l. Phonologically conservative languages (conservative in the sense that they do not show significant vowel changes) suggest that the first vowel was *a; and it is clear that the second consonant was PNCV *v, POc *p.

There is more of a problem deciding what the second vowel was, since many languages have lost it in some or all numerals or show vocalic changes. An initial approach to this problem was to examine only those languages in which the form remained constant from 6–9 (or 7–9 if 6 is aberrant). These languages are listed in regular font in Table 2. Languages in italics in that table are those which have a regular prefix-final vowel in all numerals except 9, where dissimilation occurs because of two occurrences of *v in adjacent syllables (e.g. Marino underlying /leva-vati/ > *la-vati*).

Table 2: Final vowels in the ligature *lavV-* / *ravV-*

VOWEL	LANGUAGES
-i	Banks: Vera'a
-ia	Banks: <i>Merlav</i>
-e	Banks: Lehali, Lehalurup, Nume, Dorig. Santo: <i>Tolomako, Navut, Shark Bay</i>
-ea	Banks: Mota
-a	Santo: Akei, <i>Marino</i>
-aʔa	Santo: Tasiriki
-u	Santo: <i>Malmariv</i>

François (2005) has carried out a detailed study of the reflexes of the POc vowels in Banks languages. Most of these languages have undergone significant sound changes, with only Mota being conservative. The sources of the prefix-final vowel in some of the Banks languages (cf. François 2005:490–491) are possibly as follows:

- (11) Lehali *e*: *i(C)e, *i(C)a, *i(C)o, *e(C)i, *e(C)u
 Lehalurup *e*: *e(C)i, *e(C)u
 Nume *e*: *e(C)e, *e(C)a, *e(C)o
 Dorig *e*: *e(c)e, *e(C)a, *e(C)o
 Mota *ea*: *e(C)a

This suggests to me that the original form may have been *lave(C)-a,¹⁴ i.e. a transitive verb with an object suffix, which correlates with Crowley's suggestion regarding the Paamese ligature *lahi*- deriving from 'carry'.

Now Clark (in press) has reconstructed two PNCV verbs which may be variants of the same original form: *lavi 'carry, take' and *la-i 'take, give'. It seems possible that this linker *lave-a is related to, or is an aberrant form of, *lavi-a 'carry/take-3SG'. I assume that the Tolomako-type discussed in §3 was probably the original: that is, the forms for 7–9 were *lima-lave-a-rua/tolu/vat(i), while the form for 6 was either simply *lima-lave-a or *lima-lave-a-ta (or some other form for 1). The meaning would have been something like '5 carry 1/2/3/4'. In most languages with this ligature, there was a certain amount of redundancy in the trimorphemic form, and this generally reduced to a bimorphemic form, with the numeral 1 being dropped from 6 and *lima 5 being dropped from 7–9.

An apparent reflex is found outside Vanuatu in at least one language: Tolai (Kuanua) of East New Britain has *lap-tikai*, *lavu-rua*, *lavu-tul* and *lavu-wat* for the numerals 6–9 (compare *tikai*, *au-rua*, *au-tul*, *ai-wat* for 1–4). Lynch, Ross and Crowley (2002:72) proposed POc *(l,r)apo- to account for this correspondence, but it is likely that this resemblance is accidental, and that the Tolai prefix may derive from *lap* ~ *lav* 'follow'.

4.2 *[la]kau-

There is a number of languages which have a ligature which is, or contains, *kV-. Some of these were listed in (5) in the discussion in §2.4 of 'incipient innovative systems' in south Santo/north Malakula. An apparently related form is the ligature *ha*- in Tasiriki *ha-ʔese* 6 (see §3 above).¹⁵ Seemingly similar forms are found in southeast Malakula (Aulua, Port Sandwich and Avok below) and Epi (Lewo in (2), where *k > Ø, and Bonkovia and Mae-Morae below), most with a preceding syllable.

(12)	Aulua	Pt. Sandwich	Avok	Bonkovia	Mae-Morae
6	<i>dro-vokhol</i>	<i>emut-sukai</i>	<i>ŋə-mekh-cəkai</i>	<i>wo-ra</i>	<i>lok-rogai</i>
7	<i>drokhu-rue</i>	<i>e-mokh-ii</i>	<i>ŋə-mekhu-ru</i>	<i>oko-lua</i>	<i>loku-lua</i>
8	<i>drokh-til</i>	<i>e-mokhu-röj</i>	<i>ŋə-mekho-rær</i>	<i>oko-rolu</i>	<i>lok-rol</i>
9	<i>drokh-ves</i>	<i>e-mokhu-pac</i>	<i>mekho-pæc</i>	<i>oko-veri</i>	<i>lak-var</i>

These forms suggest *ko- or *ku-, with a preceding element. In Aulua and a number of Epi languages (like Mae-Morae in (12)), the preceding element is a liquid + vowel, and I suggest here that the original form of this ligature may be the PNCV verb *lakau (~*lakawa) 'cross over' (Clark 2009). This makes sense semantically: having counted the fingers of one hand, one then 'crosses over' to the other hand to begin counting at 6. This is semantically (though not phonologically) identical to the ligature *benā* in Banoni and

¹⁴ Pawley (1972:47) reconstructed this form as *lapu-.

¹⁵ The regular reflex of *k in Tasiriki is ʔ, but *kV seems to be reflected as *hVʔ*, cf. 'Proto-Santo' *ka-kara 'red' (Wusi *kara*, Nokuku *kekara*, Tolomako *ɣavara*) > Tasiriki *haʔara*.

Piva in Bougainville (Lincoln in press): *benā* ‘cross over’ is 6, and *benā-2/3* are 7–8 (9 is *visa*, possibly a metathesised form of **siwa*).

The ligature **lakau-* presumably reduced to **kau-* in a number of languages. Note its phonological similarity to the forms in the ‘incipient innovative’ languages given in §2.4 (Araki *hai-*, Nese *kho-*, *khe-*, Botovro *ho-*, *khe-*, Tirax *khɔ-*, *khɛ-*, Malua Bay *kha-*, *kho-* and Vovo *khe-*, *kho-*).

4.3 **zau-*

There is a group of languages in Malakula which have a ligature whose form was possibly **zau-*, with initial PNCV **z* < POC **j*, (although the vocalic element undergoes some variation in some of these languages). Some examples:¹⁶

(13)	Avava	Neve’ei	Naman	Neverver	V’änen Taut	Tape	Nāti	Nahavaq
6	<i>sout</i>	<i>nsouh</i>	<i>nsous</i>	<i>jos</i>	—	—	<i>seu-siʔ</i>	<i>sow-siʔ</i>
7 / _* <i>rua</i>	<i>sou-</i>	<i>nsu-</i>	<i>nsu-</i>	<i>jo-</i>	<i>sa-</i>	<i>ji-</i>	<i>seu-</i>	<i>sow-</i>
8 / _* <i>tolu</i>	<i>se-</i>	<i>nsu-</i>	<i>nsu-</i>	<i>jo-</i>	<i>sa-</i>	<i>ji-</i>	<i>seu-</i>	<i>sow-</i>
9 / _* <i>pat(i)</i>	<i>sa-</i>	<i>nsa-</i>	<i>nso-</i>	<i>jo-</i>	<i>sa-</i>	<i>ji-</i>	<i>seu-</i>	<i>sow-</i>

A possible origin for this form—though I am not at all confident about this— may be PNCV **sabo*. Clark (in press) glosses this as ‘ignorant, incompetent, lost’, though in some languages the meaning has shifted to ‘other, different’, as in Paamese *savo* ‘different’, *mee-savo* ‘someone else, stranger’, Naman *i/nsəb* ‘other, different’; so 7 would be ‘other (hand)-2’. It is also not clear whether this form is related in any way to the ligature found in two Erromangan languages given in (9): Ura *sine-*, *sini-* and Utaha *simna-*, *simni-* (and note Ura *sai* ‘other, different’).

5 Discussion

Two different general types of innovative systems seem to have developed in the Vanuatu–New Caledonia area—imperfect decimal and true quinary systems—with a mixed system occurring in the central area. The basic forms of relevant numerals in these systems are given in Table 3.

Table 3: General types of innovative systems

	Imperfect decimal			Mixed	Quinary
	* <i>lave-a-</i>	*[<i>la</i>]kau-	* <i>zau-</i>		
6	5-carry(-1)	cross.over-1	other(-1)	{ <i>same</i> <i>as</i> <i>imperfect</i> <i>decimal</i>	5(-and)-1
7	(5-)carry-2	cross.over-2	other-2		5(-and)-2
8	(5-)carry-3	cross.over-3	other-3		5(-and)-3
9	(5-)carry-4	cross.over-4	other-4		5(-and)-4
10	ten			two fives/hands	5(-and)-5, two fives/hands
20	two tens			two tens	one person

¹⁶ V’änen Taut and Tape do not use this form in the numeral 6, which is 5–1. In Nāti and Nahavaq, the form for 6 is transparently LIG-1, but this is not the case in the other languages, and the form for 6 is given in full in the first four languages in (13).

5.1 Imperfect decimal, quinary and mixed systems

The imperfect decimal type is found throughout north and central Vanuatu (except for those languages with pure decimal systems) as far south as the middle of Ambrym. It is in fact basically a decimal system, with its only deviation being that it has replaced the terms for 6–9 with compounds, but compounds of a quite different type from those found in the far south. Many of these languages also have uncompounded forms for 100 and 1000. (Indeed, there are even relics of original terms for 6–9 in some of these languages. In an intriguing footnote to his discussion of the quinary Lewo numeral system, illustrated in (2) above, Early (1994:213) says:

The Lewo counting system for spirit-creatures is still widely known, retained as an item of conscious cultural knowledge, but without any current function. The numbers, which go from 1 to 10, are: *taka, luaka, telka, verka, limka, kona, isi, varo, siwe, kuru*.

The numerals 6–9 reflect POc *onom, *pitu, *walu and *siwa.)

True quinary systems are found to the south of this area. In these systems, numerals effectively stop at five, and everything above that is an additive compound until twenty is reached, which is usually a compound meaning ‘one person’—i.e. all fingers and toes have been counted. Addition then starts again from 21–40 (‘two people’), 41–60 (‘three people’), etc. There are no monomorphemic words for 100 or 1000. Mixed systems occur in parts of central Vanuatu (Epi, Shepherds and Efate and Erromango).

Given that true or complete decimal systems still survive in parts of northern Vanuatu, and given the distribution of the imperfect decimal and quinary types within Vanuatu and New Caledonia, the following sequence of events suggests itself to me:

- I. The original settlers of the area (somewhere in northern Vanuatu) had a true decimal system.
- II. The first innovation was to change to an imperfect decimal system, using some variant of LIG-1–4 for 6–9, but otherwise leaving the decimal system unchanged.
- III. A true quinary system developed out of this, somewhere in the centre or south of Vanuatu, and the two systems were in some kind of competition for some time, which resulted in the mixed systems in the central area.

Proposal I is uncontroversial: the full decimal system is retained in at least some Vanuatu languages, which are geographically non-contiguous (including the Lewo counting system for spirit-creatures).

Proposal II is also uncontroversial. Settlement of Vanuatu was generally in a north-to-south direction. The imperfect decimal system is widespread in the north, and elements of it show up as far south as Anejoñ in the extreme south. The use of the LIG-1–4 system probably evolved from counting on the fingers of one hand and then crossing over to the other (cf. also Lincoln in press); the bare use of the ligature here (rather than 5-LIG-1–4), along with the fact that the numeral 6 is often different in form from 7–9 and indicates a ‘crossing over’, tend to support this. The fact that some of these languages preserve the ligature *dum^wa- for 11–19, found also in languages with true decimal systems, along with words for 100 and 1000 (which may, of course, have come to mean ‘fairly high number’ and ‘somewhat higher number still’), lends support to this hypothesis.

Proposal III recognises that quinary systems have a restricted geographical distribution, with no evidence of any being found anywhere north of Southeast Ambrym.

5.2 On the origins of the systems

As I mentioned at the beginning of this paper, Blust (2005) has suggested that the ‘quinary’ systems—by which I assume he means innovative (in my terminology), not just quinary—in this area arose out of Papuan contact *in situ*. That suggestion, of course, is impossible to disprove, since it implies also that any Papuan language or languages spoken in Vanuatu–New Caledonia which influenced the Oceanic languages of this region have since disappeared.

The nearest extant Papuan languages are in the central Solomon Islands. All Oceanic languages of the Solomon Islands (excluding Bougainville) as far south as Makira have pure decimal systems. Michael Dunn (pers. comm. 3 October 2007) says of the Solomons Papuan languages that:

there aren’t any quinary systems there any more. All the Solomons Papuan languages have productive decimal systems, but there’s internal evidence ... that they developed out of a quinary system.

There are, however, quinary systems close to northern Vanuatu, in the languages of the Temotu Province of Solomon Islands, all of which are now known to be Oceanic and not Papuan (Lincoln 1978, Ross and Næss 2007). There are nine languages in this area: all three in Vanikoro and one in Utupua have straightforward decimal systems. Of the remainder, Nagu in Santa Cruz has a subtractive system (7 = LIG-3, 8 = LIG-2, 9 = LIG-1); and Äiwoo in the Reef Islands, Natügu in Santa Cruz and two languages in Utupua have innovative systems of the same general type as northern Vanuatu (i.e. 7 = LIG-2), though Natügu in (14) appears to have 2-LIG. The relevant numerals in these languages are listed in (14) (1–9 from Ross and Næss 2007, 10 from Tryon and Hackman 1983).

These languages also appear to have forms for 100 and 1000 which are not compounds (Tryon and Hackman 1983). In other words, with the exception of the subtractive Nagu, these are of a similar type to the imperfect decimal systems in much of Vanuatu, though there do not appear to be any cognates among the ligatures.

Lean (1992) conducted a thorough study of Papuan (and Oceanic) counting systems on the island of New Guinea and neighbouring islands. His conclusions as to the distribution of various kinds of systems are set out in Table 4 (Table 63 in the original). While recent research has suggested that the East Papuan phylum is no longer a valid grouping (Ross 2001), I leave that heading in the table (with that column shaded) since it encompasses the extant Papuan languages geographically closest to Vanuatu and New Caledonia. The table also conflates some types which differ in only very minor ways, and uses different names for some of the types of systems to conform with the usage in this paper. The two innovative systems found in Vanuatu–New Caledonia are also shaded on this table.

(14)	Äiwoo	Natügu	Nagu	Nebao	Tanibili
	Reef Is.	Santa Cruz	Santa Cruz	Utupua	Utupua
1	<i>nyigi</i>	<i>tʌe-sʌ</i>	<i>täte/ëte</i>	<i>tua</i>	<i>suo</i>
2	<i>li-lu</i>	<i>li</i>	<i>la-li</i>	<i>l-lu</i>	<i>bu-yu</i>
3	<i>eve</i>	<i>tii</i>	<i>lʌ-tii</i>	<i>tʰɔ</i>	<i>bo-kʷo</i>
4	<i>u-vä</i>	<i>pʷä</i>	<i>lʌ-fɔ</i>	<i>hia</i>	<i>mā-piɔ</i>
5	<i>vi-li</i>	<i>nʌlvü</i>	<i>lʌ-mëf</i>	<i>haji</i>	<i>kavili</i>
6	<i>pole-gi</i>	<i>e-sʌ-mʌ</i>	<i>lʌ-mëtʰemë</i>	<i>uru</i>	<i>kavili suo</i>
7	<i>pole-lu</i>	<i>ë-li-mʌ</i>	<i>tumë-tu</i>	<i>va-lu</i>	<i>suo-vi-yo</i>
8	<i>pole-e</i>	<i>ë-tii-mʌ</i>	<i>tumë-li</i>	<i>va-ro</i>	<i>ve-vi-ro</i>
9	<i>polo-uvä</i>	<i>ë-pʷä-mʌ</i>	<i>tumë-te</i>	<i>wa-hia</i>	<i>vere-ve-pio</i>
10	<i>nugolu</i>	<i>nʌpnu</i>	<i>nëpnu</i>	<i>ɲalɔ</i>	<i>vere-ɲalu</i>

Table 4: Distribution of Papuan numeral systems

Types	West Papuan	Torricelli	Sepik-Ramu	Trans New Guinea	East Papuan	Minor Phyla	Total
Binary	0	0	3	39	0	0	42
Binary + 5	0	24	13	134	2	3	176
Quinary	0	2	17	52	1	7	79
Base-4	0	0	1	6	0	2	9
Base-6	0	0	0	5	0	0	5
Body parts	0	0	8	58	0	4?	70?
Imperfect decimal	7	0	3	8	12	3	35
Decimal	3	0	1	3	8	0	16

Lean (1992, §3.4) notes that the imperfect decimal systems of the ‘East Papuan phylum’ are mainly located in New Britain and Bougainville. But Lincoln (in press) has clearly demonstrated that the imperfect decimal systems of Banoni and Piva (and, probably, other Bougainville Oceanic languages) have *not* developed as a result of Papuan influence, but as a result of the physical nature of the counting process. I believe I have demonstrated that the same is true of the imperfect decimal systems in Vanuatu–New Caledonia. The only point at issue, then, is how the quinary systems developed; these differ from imperfect decimal systems mainly in the loss of forms for 10 and in using the person-type construction rather than numerals for 20. Was this due to contact? If not, and this seems highly unlikely given their location, was it simply due to the ‘regularisation’ of LIG-1–4 systems as 5 + 1–4 systems? Proposal III above also referred to the two systems being in ‘competition’ for some time. Just as Araki and the Malakula languages discussed in §2.4 seem to be showing a change from decimal to imperfect decimal, so too the data in (14) suggest a change in progress from imperfect decimal to quinary; but in both cases the change is incomplete. The ligatures (*lave-a and the others) lost their transparency as ‘carrying’ or ‘cross over’ verbs, partly due to phonological change and attrition, and were replaced by a more transparent compound 5 + 1–4.

6 Conclusion

In this paper, I have described in some detail the numeral systems of the languages of Vanuatu and New Caledonia, and shown how they developed morphologically out of the POC decimal system which is still found in some northern Vanuatu languages. I have not attempted to account for the origin of these systems, or for the motivation which led to a change from decimal to innovative systems, but have provided adequate data so that the Papuan contact hypothesis might be tested. When it is tested, however, it is worth noting that the greatest deviation from the standard Oceanic decimal system comes in those languages which are geographically most distant from known Papuan-speaking areas.

Appendix: Orthography and sources of data

Standard orthography is used in most cases (but not for languages which do not have one), though the velar nasal and glottal stop are consistently represented as *ŋ* and *ʔ* and the schwa in Vanuatu languages as *ə*. The symbols *m'*, *v'* and *p'* in various languages are apicolabials; labiovelars (often marked by tildes in some Vanuatu languages) are consistently written with a superscript *w*; *kh* in Malakula languages represents a velar fricative (variably voiced depending on position), irrespective of what the standard orthography uses; Araki *d* is a flap, *r* is a trill, whereas in Nese and Uripiv *r* is a flap and *rr* a trill.

Specific data sources for many of the languages discussed in this paper are given below where there is no specific reference in the text. Data for those which do not appear on this list, and additional data for some which do, come from either Tryon (1976), Charpentier (1982), Lynch, Ross and Crowley (2002), Clark (in press) and/or the *Austronesian Basic Vocabulary Database* (Greenhill, Blust and Gray 2005–09). Data for Nahavaq, Neverver, Tirax and Uripiv come from unpublished wordlists/dictionaries compiled by Laura Dimock, Julie Barbour, Amanda Brotchie and Ross McKerras respectively, while Nengone data are from Claire Moyse.

Sources for other languages are: Anejoñ (Lynch 2000); Araki (François 2002); Avava (Crowley 2006a); Drehu (Moyse-Faurie 1983); Iaai (Ozanne-Rivierre 1976); Lenakel (Lynch 1978); Lewo (Early 1994); Merei (Chung 2005); Mota (Codrington 1885); Mwotlap (François in prep.); Naman (Crowley 2006b); Nāti (Crowley 1998b); Nemi (Haudricourt and Ozanne-Rivierre 1982); Nese (Crowley 2006c); Paamese (Crowley 1982); Port Sandwich (Charpentier 1979); South Efate (Thieberger 2006); Southeast Ambrym (Parker 1970); Sye (Crowley 1998a); Tape (Crowley 2006d); Ura (Crowley 1999); Utaha (Lynch 2001); V'ënen Taut (Fox 1979); Xârâcùù (Moyse-Faurie 1995).

References

- Blust, Robert. 2005. Review of John Lynch, Malcolm Ross and Terry Crowley, *The Oceanic languages*. Richmond, UK: Curzon Press, 2002. *Oceanic Linguistics* 44.544–558.
- Charpentier, Jean-Michel. 1979. *La langue de Port-Sandwich (Nouvelles-Hébrides): Introduction phonologique et grammairale*. Paris: SELAF.
- . 1982. *Atlas linguistique du Sud-Malakula/Linguistic atlas of South Malakula (Vanuatu)*. Paris: CNRS/L'Agence de Coopération Culturelle et Technique. [Langues et cultures du Pacifique 2.]

- Chung, Yun Shing Anthony. 2005. *A descriptive grammar of Merei (Vanuatu)*. Canberra: Pacific Linguistics.
- Clark, Ross. 1985. Languages of north and central Vanuatu: groups, chains, clusters and waves. In Andrew Pawley and Lois Carrington, eds *Austronesian linguistics at the 15th Pacific Science Congress*, 199–236. Canberra: Pacific Linguistics.
- 2009. **Leo Tuai: a comparative lexical study of North and Central Vanuatu languages*. Canberra: Pacific Linguistics.
- Codrington, R.H. 1885. *The Melanesian languages*. Oxford: Clarendon Press.
- Comrie, Bernard. 2008. Numeral bases. In Martin Haspelmath, Matthew S. Dryer, David Gil and Bernard Comrie, eds *The world atlas of language structures online*, cht. 131. Munich: Max Planck Digital Library. <http://wals.info/feature/131>.
- Crowley, Terry. 1982. *A dictionary of Paamese*. Canberra: Pacific Linguistics.
- 1998a. *An Erromangan (Sye) grammar*. Oceanic Linguistics Special Publication No.27. Honolulu: University of Hawai'i Press.
- 1998b. A salvage sketch of Nāti (Southwest Malakula, Vanuatu). In Darrell Tryon, ed. *Papers in Austronesian linguistics* No.5, 101–148. Canberra: Pacific Linguistics.
- 1999. *Ura: a disappearing language of southern Vanuatu*. Canberra: Pacific Linguistics.
- 2006a. *The Avava language of central Malakula (Vanuatu)* [ed. by John Lynch]. Canberra: Pacific Linguistics.
- 2006b. *Naman: a vanishing language of Malakula (Vanuatu)* [ed. by John Lynch]. Canberra: Pacific Linguistics.
- 2006c. *Nese: a diminishing speech variety of Northwest Malakula* [ed. by John Lynch]. Canberra: Pacific Linguistics.
- 2006d. *Tape: a declining language of Malakula (Vanuatu)* [ed. by John Lynch]. Canberra: Pacific Linguistics.
- Early, Robert. 1994. A grammar of Lewo, Vanuatu. PhD thesis. Canberra: The Australian National University.
- Fox, G.J. 1979. *Big Nambas grammar*. Canberra: Pacific Linguistics.
- François, Alexandre. 2002. *Araki: a disappearing language of Vanuatu*. Canberra: Pacific Linguistics
- 2005. Unraveling the history of the vowels of seventeen Northern Vanuatu languages. *Oceanic Linguistics* 44.443–504.
- In prep. Dictionnaire mwotlap–français. Electronic manuscript. Paris: LaCITO, CNRS.
- Greenhill, Simon, Robert Blust and Russell Gray. 2005–09. *The Austronesian Basic Vocabulary Database*. <http://language.psy.auckland.ac.nz/austronesian>.
- Haudricourt, André-Georges and Françoise Ozanne-Rivierre. 1982. *Dictionnaire thématique des langues de la région de Hienghène (Nouvelle-Calédonie): Pije, fwâi, nemî, jawe*. Paris: SELAF.

- Lean, Glendon Angove. 1992. Counting systems of Papua New Guinea and Oceania. PhD thesis. Lae: Papua New Guinea University of Technology.
<http://www.uog.ac.pg/glec/thesis/thesis.htm>
- Lincoln, Peter C. 1978. Reef-Santa Cruz as Austronesian. In Lois Carrington and S.A. Wurm, eds *Second International Conference on Austronesian Linguistics: Proceedings*, 929–67. Canberra: Pacific Linguistics.
- In press. Count not on substrata. In John Bowden and Nikolaus Himmelmann, eds *A journey through Austronesian and Papuan linguistic and cultural space: papers in honour of Andrew K. Pawley*. Canberra: Pacific Linguistics.
- Lynch, John. 1978. *A grammar of Lenakel*. Canberra: Pacific Linguistics.
- 2000. *A grammar of Anejom̃*. Canberra: Pacific Linguistics.
- 2001. *The linguistic history of Southern Vanuatu*. Canberra: Pacific Linguistics.
- 2007. The Malakula linkage of central Vanuatu. Paper presented to the Seventh International Conference on Oceanic Linguistics, Nouméa, July 2–6.
- Lynch, John, Malcolm Ross, and Terry Crowley. 2002. *The Oceanic languages*. Richmond, UK: Curzon Press.
- Moyse-Faurie, Claire. 1983. *Le drehu, langue de Lifou (Iles Loyauté)*. Paris: SELAF.
- 1995. *Le xârâcùù, langue de Thio-Canala (Nouvelle-Calédonie): Éléments de syntaxe*. Paris: Peeters.
- Ozanne-Rivierre, Françoise. 1976. *Le iaai, langue mélanésienne d'Ouvéa (Nouvelle-Calédonie): Phonologie, morphologie, esquisse syntaxique*. Paris: SELAF.
- Parker, Gary J. 1970. *Southeast Ambrym dictionary*. Canberra: Pacific 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, 1–142. Honolulu: Bernice Pauahi Bishop Museum.
- Ross, Malcolm. 2001. Is there an East Papuan phylum? Evidence from pronouns. In Andrew Pawley, Malcolm Ross and Darrell Tryon, eds *The boy from Bundaberg: Studies in Melanesian linguistics in honour of Tom Dutton*, 301–21. Canberra: Pacific Linguistics.
- Ross, Malcolm and Åshild Næss. 2007. An Oceanic origin for Äiwoo, the language of the Reef Islands? *Oceanic Linguistics* 46(2).456–498.
- Thieberger, Nicholas. 2006. *A grammar of South Efate: An Oceanic language of Vanuatu*. Honolulu: University of Hawai'i Press. [Oceanic Linguistics Special Publication No.33.]
- Tryon, Darrell. 1973. Linguistic subgrouping in the New Hebrides: A preliminary approach. *Oceanic Linguistics* 12.303–51.
- 1976. *New Hebrides languages: an internal classification*. Canberra: Pacific Linguistics.
- Tryon, D.T. and B.D. Hackman. 1983. *Solomon Islands languages: an internal classification*. Canberra: Pacific Linguistics.