Journal of the Linguistic Society of Papua New Guinea

ISSN: 0023-1959

Vol. 34 No. 1, 2016

Numeral Systems, Internal Subgrouping, and Language Contact in Malakula

John Lynch
UNIVERSITY OF THE SOUTH PACIFIC
johnlynch123@gmail.com

This paper examines developments in the numeral systems of the 32 languages of Malakula, central Vanuatu, illustrating the complex linguistic history of this small but linguistically complex island. I will try to assess whether subgrouping hypotheses made on the basis of phonological evidence alone can be confirmed or further refined. At the same time, I will suggest that certain of these developments may well have spread through contact.

1. INTRODUCTION. In a recent paper (Lynch to appear), I proposed a tentative subgrouping of the 32 languages of Malakula (central Vanuatu). This was based solely on phonological innovations, since there is very little grammatical information available on perhaps two-thirds of these languages.

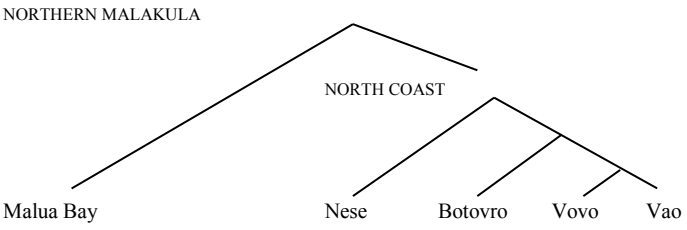
One area of morphology—or at least historical morphology—where the data are adequate to permit detailed study across all 32 languages, however, is numerals and counting systems. This paper is an attempt to test that phonologically-based subgrouping against subgrouping hypotheses suggested by an examination of these numeral systems, and to examine which developments from Proto-Oceanic (POC) may have taken place through contact rather than through inheritance.

2. MALAKULA: BACKGROUND

2.1 SUBGROUPING. There are 32 languages currently spoken on Malakula, though a few of these are moribund (and there were also some other languages that have died out altogether). There is some evidence that all the languages of Malakula belong to a single linkage[1] within the Central Vanuatu subgroup of Southern Oceanic; although that evidence is not strong, there is no evidence that suggests that different Malakula languages belong to different subgroups of Central Vanuatu.

A recent study (Lynch to appear) proposed an internal subgrouping for these languages, based on shared phonological innovations (including irregular phonological developments in particular lexical items). It suggested that the Malakula linkage consists of three higher-order component groups: a Northern Malakula subgroup, an Eastern Malakula linkage, and a Western Malakula linkage. The location of the languages and subgroups can be found on the map.

FIGURE 1. THE NORTHERN MALAKULA SUBGROUP



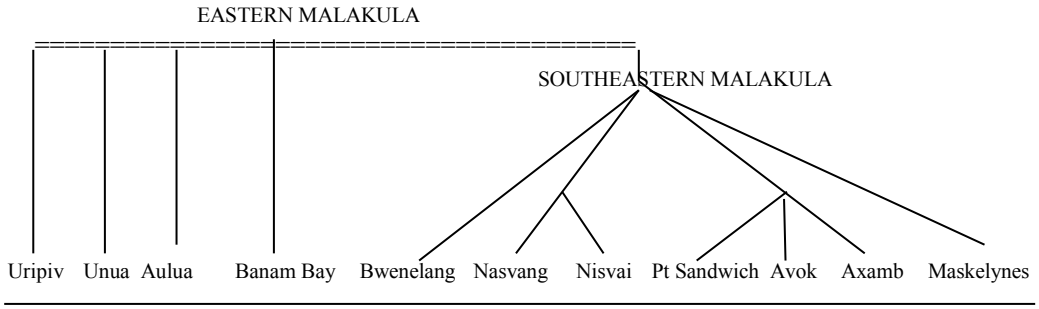
The following comments on the strength of the various hypotheses outlined in Lynch (to appear) are relevant to the discussion on numerals to follow.

a. The Northern Malakula subgroup (see figure 1, is marked by two innovations: POC *i > *u* / _ Co, and the merger of *d, *dr, and *r.[2]
b. Within Northern Malakula, languages of the North Coast subgroup share a complex set of vowel loss rules whereby final vowels were not lost if the penult was high and the ultima nonhigh (thus *na-kutu ‘louse’ > Nese *naxut* but *na-puko ‘morning’ > *nev’xe*); and, if a final consonant was retained, a paragogic vowel was added after that consonant if the final vowel or the root was high (thus *na-p^{w}ilak ‘lightning’ > Nese *nev’ilax* but *ñamuk ‘mosquito’ > *namxo*).
c. The Eastern Malakula linkage is illustrated in figure 2, where double lines represent linkages. All languages share the merger of *d and *dr, as distinct from *r, but this is also found in some languages outside the linkage. Eight of the eleven Eastern languages share the unexpected development of*t as *j (rather than the

[1] A *linkage* derives from an ancestral dialect chain, with different innovations distributed differently over the membership, and no single innovation shared by all members. This contrasts with a *subgroup*, whose members all derive from a single ancestral language and share some innovations.

[2] Many other Malakula languages merge two of these three protophonemes, but only the Northern languages merge all three.

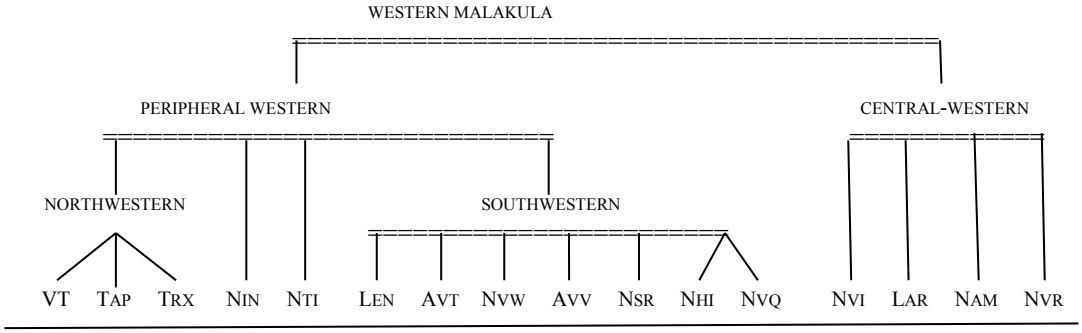
FIGURE 2. THE EASTERN MALAKULA LINKAGE



expected voiceless *s or subsequent *h) when before *i or *e: thus while *mate 'die, dead' has voiceless reflexes of *t in non-Eastern languages (thus Tape *mes*, Neverver *mas*), we find instead forms with voiced palatalization like Uripiv *e-mij*, Unua *mej*, or forms that derive from voiced *j rather than voiceless *s, like Port Sandwich *mac*, in Eastern languages.

d. Within the Southeastern sublinkage, five languages show *t > default *r*, and five or six show merger of *l and *r with subsequent complex splitting in some of these. There are also a few irregular phonological developments in lexical items (like *quti 'penis' > **qitu).
e. The Western Malakula linkage consists of two large sublinkages, Peripheral and Central-Western, plus two languages—Ninde and Nāti—which seem to be intermediate between the two as far as their subgrouping affiliations are concerned. The makeup of this linkage is in figure 3.[3]

FIGURE 3. THE WESTERN MALAKULA LINKAGE



f. The Central-Western linkage is only very weakly defined; there are no phonological innovations per se, but there are a few irregular lexical developments (like *bisu- 'fingernail' > **sibu-). Within the Western Malakula linkage, the distinction between Peripheral and Central-Western languages seems reasonably clear.
g. As to the Northwestern sublineage, V'ënen Taut shares loss of *p before *u with Tape (but also Ninde) and merger of *s and *j only before *i with Tirax (and also Nāti).
h. A number of the Southwestern sublinkage languages show complete merger of *s and *j, and all show a number of phonological irregularities in certain lexical items (like *sulati 'worm' > **dulati).
i. Ninde and Nāti have links with both Northwestern and Southwestern groupings, and their position is ambivalent; while Tirax remains an anomaly, showing little in common with any other languages, though it *seems* to fit with the Northwestern group.

It is with this background that I approach the examination of the numeral systems.

[3] Language names in figure 3 are abbreviated for reasons of space. Reading left to right, these abbreviations are: VT, V'ënen Taut; TAP, Tape; TRX, Tirax; NIN, Ninde; NTI, Nāti; LEN, Lendamboi; AVT, Aveteian; NVW, Navwien; AVV, Avava; NSR, Nasarian; NHI, Naha'ai; NVQ, Nahavaq; NVI, Neve'ei; LAR, Larëvat; NAM, Naman.

MAP. MALAKULA LANGUAGES AND SUBGROUPS

2.2 DATA AND DATA SOURCES. Data sources for the numerals in the Malakula languages are as outlined in the appendix. Where a name is used without a following date, this indicates that that person kindly supplied unpublished data for that language.

Tryon (1976) and Charpentier (1982) are large-scale surveys, with all the limitations (potential for incorrect elicitation, phonemic uncertainty, etc.) that that implies. Where there is a disagreement between the two Charpentier sources, I follow the 1987 paper for the form of the numeral; however, some of Charpentier's proposals in that paper as to historical developments are rather fanciful, and I largely ignore them.

I use a common orthography in citing data so as to facilitate cross-language comparison. For example, the velar fricative is variously written *x*, *h*, and *kh* in the orthographies of different languages, but here I use *x* for all languages. Voiced stops are prenasalized in all languages, while apicolabials (or linguolabials) are written with a following ': thus *v'*, *b'*, *m'*. Other symbols requiring explanation are *rr*, a trill contrasting with a flap *r* in some languages; *dr*, a voiced prenasalized alveolar trill; and *bb*, a voiced bilabial trill.

Numerals in Malakula languages tend to be verbs. In many languages, their citation form includes what is, or was, a 3SG prefix (most usually *i-*). In some languages, they vary for mood, as in Neverver (Barbour 2012:137): contrast realis-marked *i-ru* '2' and *i-tl* '3', used in realis clauses, with irrealis-marked *ib-ru* and *ibi-tl*, used in irrealis clauses.

3. NUMERAL SYSTEMS IN SOUTHERN OCEANIC. Proto-Oceanic had a decimal system, with the roots 1–9 being essentially monomorphemic, and 10, 20, and the like being '1×10', '2×10' (with a single morpheme representing '10'). The forms were probably as follows:

(1)	1	*ta-sa, *sa-kai, *tai, *kai	6	*onom
	2	*rua	7	*pitu
	3	*tolu	8	*walu
	4	*pat, *pati	9	*siwa
	5	*lima	10	*sa-ŋa-puluq (1-ligature-10)

There seems to have been a counting prefix *ka- used at least with the numerals 1–9: *ka-tai, *ka-rua, *ka-tolu, … 'one, two, three, ...'. Decades are based on the root *puluq 'ten' preceded by the ligature *ŋa: thus *rua-ŋa-puluq '20', *tolu-ŋa-puluq '30', and so on. Intervening numerals were linked to the decade marker by *ma 'and': thus *rua-ŋa-puluq ma tolu '23'. There was a term for '100', *Ratu(s). See Lynch, Ross, and Crowley (2002:72–74) for further discussion.

In Lynch (2009), I outlined the development of the numeral system in the Southern Oceanic subgroup, which consists of the 120 or so non-Polynesian languages of Vanuatu and New Caledonia. While some of these languages retain the decimal system virtually unchanged, many others show reductions of various kinds. Three other kinds of systems occur with some frequency:

- IMPERFECT DECIMAL SYSTEMS: these differ from decimal systems only in that the numerals 7–9 are compounds generally involving the numerals 2–4 in some way; 6 may also follow this pattern, or may be constructed irregularly. In these systems, 10 and 20 are constructed in the same way as for decimal systems.
- QUINARY SYSTEMS: 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'.
- MIXED SYSTEMS: these have features of both imperfect decimal and quinary systems. Typically, 10 is 'two fives/hands', but 20 is 'two tens'.

An example of each type of system is given in (2), with the reflex of the POC root separated from other material by a slash:

(2)	**Imperfect decimal**	**Quinary**	**Mixed**
	MEREI (Santo)	LENAKEL (Tanna)	SYE (Erromango)
1	ese	ka/rena	hai/(teven)
2	ruwa	k/iu	ndu/ru
3	tolu	kə/sil	nde/hel
4	vat	ku/vər	nd/vat
5	lima	kati/lum	suk/rim
6	ma/ravo	katilum-karena	me/hikai
7	ravo/rua	katilum-kiu	sukrim-nduru
8	rap/tol	katilum-kəsil	sukrim-ndehel
9	rait/at	katilum-kuvər	sukrim-e-ndvat
10	saŋavul	katilum-katilum	na/rwolem
20	ŋavul-rua	ieramim karena rəka	narwolem nduru

Merei shows a feature that is quite common: while many languages use '5-ligature-2', say, for '7', many others (like Merei) use simply 'ligature-2'.

In that 2009 paper, I suggested that, while closely related languages may have different numeral systems, there are some geographical tendencies. Languages north of Epi in central Vanuatu have either decimal or imperfect decimal systems; in the central area of Southern Oceanic, languages of Epi, Efate, and Erromango have mixed systems; while in the south, languages of Tanna, Aneityum, and New Caledonia have quinary systems.

4. DEVELOPMENTS IN MALAKULA NUMERAL SYSTEMS. As to what was known of the numeral systems of Malakula languages at the time Lynch (2019) was published, seven languages in the far north of the island were classified as having a decimal system (and these are the southernmost pure decimal languages within Sothern Oceanic), and the remainder as having imperfect decimal systems. Since then, however, more data have become available on Malakula languages, and more work has been done on their internal relationships, and it appears that there are a number of complications that I was not aware of at the time. It seems appropriate, therefore, to examine the development of numeral systems in some detail to see what inferences we can make about language history.

In this section, I will simply present the developments that have taken place, and comment where necessary on th relevance of each to the question of historical development (subgrouping or contact). I will leave until section 5 any wider evaluation of the subgrouping implications.

4.1 OVERALL SYSTEMIC DEVELOPMENTS

4.1.1 Imperfect decimal systems. Only seven Malakula languages retain the original decimal system unchanged: all five members of the Northern Malakula subgroup, plus one Eastern language (Uripiv),[4] and one Western language (Tirax). The bulk of the languages of the Eastern and Western linkages shifted from a decimal to an imperfect decimal system, with 6–9 being formed by additive compounds, but a monomorphemic word for 10, and 20 marked as two tens. This is exemplified in (3) by two languages from each linkage, relatively geographically and genetically distinct from each other.

[4] Strictly speaking, the "official" name of this language is Northeast Malakula (Lynch and Crowley 2001); it consists of the dialects spoken on a number of small offshore islands, including Atchin, Uripiv, Uri, Wala, and Rano. The Uripiv dialect is the most prominent, and the language as a whole is usually referred to as Uripiv.

(3)

	EASTERN		WESTERN	
	UNUA	MASKELYNES	TAPE	NEVERVER
1	soxa	e-sua	i/simək, i/sig	i-sxam
2	xe/ru	e-ru	i/ru	i-ru
3	xe/ter	i-tor	i/təl	i-tl
4	xe/vej	i-vat	i/ves	i-vas
5	xe/rim	e-rim	i/ləm	i-lim
6	morov/tes	(e-)məlev-tes	ləm/ji/s	i-jo-s
7	morov/ru	(e-)məlev-ru	ji/ru	i-jo-ru
8	morov/tor	(e-)məlev-tor	ji/təl	i-jo-tl
9	moro/pej	(e-)məla-pat	je/vet	i-jo-vas
10	saŋavur	(e-)səŋavur	i/sŋel	naŋavul
20	ŋavur xeru	səŋavur vəha-ru	i/ŋel/ru	naŋavul i-ru

Tirax is one of the two most northerly Western languages (the other being V'ënen Taut), and Uripiv the northernmost Eastern language. Two possible historical scenarios are (i) that the full decimal system was lost in an early ancestor of both the Eastern and Western linkages, except in the extreme north of both (what are now Uripiv and Tirax), or (ii) that it was completely lost in those two ancestral languages but reintroduced into Uripiv and Tirax through contact with one or more Northern languages, probably Vao, or even with Tamambo of neighboring Malo island, as a result of trade in leaf dye and mats (Huffman 1996:188). Vao appears to have been central to a number of trading networks, and Tryon (1976:81) noted also its linguistic "middleman" status in terms of his lexicostatistical classification: "It will be noted that Vao, in the Malekula Coastal sub-group, acts as a kind of link pin between the West Santo, Malekula Coastal and East New Hebrides sub-groups." I tend, therefore, to option (ii) above: that the ancestors of both the Eastern and Western linkages replaced the (perfect) decimal with an imperfect one, but that the decimal system was reintroduced into Uripiv and Tirax through contact with a Northern language, probably Vao.[5]

4.1.2 Mixed systems. There are seven languages, however, that show a departure from this imperfect decimal system. , in that the form for 20 is not a compound of '10' and '2' but is rather built on the form for 'man, person'.[6] These forms are listed below, with comments (where available) on the nature of the form:

(4) NASVANG	naməxar	= 'person'
AXAMB	ren-xavoy a-ceke-nene	= 'man-true he-1-only' (Charpentier 1987:111)
NĀTI	mwarlala	not clear, but see Nahavaq
LENDAMBOI	na-mükut	= 'person'
NAVWIEN	na-mükut	= 'person'
NAHA'AI	namaratin	*namar* 'person' + *atin* ??
NAHAVAQ	ni-morlalaʔ	*mor* 'person' + *lalaʔ* 'quiet/still' (Dimock 2009:130)

This occurs in two linkages: (i) a subset of five Southwestern languages, which are all geographically contiguous in the southwest "corner" of the island (though Ninde, which is also spoken in that area, is excluded): Lendamboi, Nāti, Nahavaq, Naha'ai, and Navwien; and (ii) two Southeastern languages, Nisvai and Axamb (though not in Nasvang, which is located between them). This suggest two things to me:

a. that the innovation developed within the southern part of the Southwestern linkage (and *may* also suggest that Nāti should be added to this grouping); and
b. that its presence across the major linkage boundary in Nisvai and Axamb could well be to due to contact with these languages to their west.

4.2 IRREGULAR PHONOLOGICAL DEVELOPMENTS OF POC ROOTS. There is considerable variation with forms for 1, which is common throughout Oceanic, and which Lynch, Ross, and Crowley (2002:73) liken to the semantic and functional variation involving English *a/an, one, single,* and *only*. The numerals 2–5 (and 6–9, where retained) are generally inherited regularly, with both versions of the POC form for 4 being found: *pat > Nese, Botovro *v'at*, Vao *xe-vat*, Tirax *vat* (*t regularly > *s* / _ *i in these languages);

[5] Note discussion of these trade networks and some aspects of linguistic borrowing in this area in Lynch and Brotchie (2010:384–86).

[6] Given the geographical distribution of these languages, one might expect Aveteian to also show it; but there is no form with that meaning in the data available to me.

and *pati > Unua *xe/vej*, Uripiv *i/vij*, Tape *i/ves*, Ninde *ves* (with regular "palatalization" of *t). There are, however, some irregularities.

4.2.1 Irregular developments in 2–4. There are three languages that have inherited the numerals 2–4 with an unexpected nasalized reflex of the initial consonant; compare the forms for 2–4 in (5) with the additive compounds 7–9 which contain the non-nasalized reflex:

(5)	AULUA	BWENELANG	NAVWIEN
2	i-**drua**	i-**dru**	i-**dru**
7	droxu-**rua**	lov-**ru**	i-drax-**ru**
3	i-**dil**	i-**dil**	i-**dül**
8	drox-**til**	lov-**təl**	i-drax-**tül**
4	i-**bos**	i-**bec**	i-**bas**
9	drox-**ves**	lo-**pe**	i-drax-**vas**

This occurs in two Southeastern languages and the (geographically distant) Navwien in the west. I cannot see contact being involved here, and we may be dealing with two (or even three) independent developments.[7]

Conversely, there is a group of Eastern languages that do the reverse with the form for 4: initial fricative (oral reflex) in the numeral 4, initial stop (nasalized reflex) in the compound 9:

(6)	**4**	**9**
UNUA	xe-**vej**	moro-**pej**
BANAM BAY	e-**vet**	ro-**pe**
NASVANG	i-**vac**	i-ləm mə-rax-**pac**
NISVAI	ŋa-**vac**	rax-**pac**
PORT SANDWICH	e-**vac**	e-mox-**pac**
MASKELYNES	i-**vat**	(e)-məla-**pat**

This is an innovation within the Eastern linkage, shared by some (but not all) Southeastern languages with Unua and Banam Bay, and which may cause us to consider adding these two languages to that sublinkage.

And there is another set of languages in which the form meaning 4 in the word for 9 has a quite distinct final phoneme from that of the form in the plain numeral 4. The languages that show this are in (7), and are grouped according to similar behavior.

(7)	**4**	**9**
BANAM BAY	e-**vet**	ro-**pe**
BWENELANG	i-**bec**	lo-**pe**
NĀTI	i-**vös**	seu-**vei**
NAHAVAQ	i-**ves**	(i)-sow-**vey**
V'ËNEN TAUT	i-**v'a**	i-sa-**v'et**
TAPE	i-**ves**	je-**vet**

The first four show the development fricative (or affricate) > glide > Ø; while the two Eastern languages are geographically close, as are the two Western languages, there is little evidence of contact between east ad west, and we are probably dealing with two independent innovations. The last two, in the closely related Northwestern languages V'ënen Taut and Tape, look as if the form in 4 reflects *pati (*ti and *s > Ø is regular in V'ënen Taut) while the form in 9 reflects *pat.

4.2.2 POC *saŋapuluq '10' > **laŋapuluq. There is an irregular development with the word for 10, occurring in Ninde, Nāti, and all the Southwestern linkage languages, and this is the totally unmotivated change of the initial *s of *saŋapuluq to *l. These forms are shown in (8), along with reflexes of *saman 'outrigger', *saqat 'bad', and *susu 'breast', which show the regular development of *s (variously *s*, *h*, or Ø) in these languages.

[7] There is often fluctuation in Malakula languages between voiceless and voiced (or oral and nasal) initial consonants in verbs (see (Lynch 2008:296–97). Since numerals are, or were, verbs, this may be the explanation here.

(8)	*saŋapuluq	*saman	*saqat	*susu
	'ten'	'outrigger'	'bad'	'breast'
NINDE	lhaŋal	ne/sep	—	n/u- [*s > Ø regular]
NĀTI	laŋavöl	ni/sem	—	ni/süsü-
AVAVA	laŋal	—	se	a/sih
NASARIAN	i-laŋavul	—	sey	na/u- [*s > Ø regular]
AVETEIAN	i-laŋavil	—	set/əgan	nen-səs-tni
LENDAMBOI	i-laŋavul	nə/sem	set	na/səs
NAVWIEN	i-laŋvül	ni/sem	sat	nə/süsü
NAHAVAQ	laŋavul	ne/sem	het	nu/huhu
NAHA'AI	i-laŋvül	ne/sem	hat	nu/süsü

This innovation could provide evidence that Ninde and Nāti belong in the Southwestern linkage. This will be evaluated later, in light of other developments in these languages.

4.2.3 When 6 is not 5 + 1. There are languages, like those in (9a), where the form for 6 is clearly (5)-LIG-1; and there are others, like those in (9b), where the form for 1 following the ligature is clearly derived from the numeral 1, though not phonologically identical.

(9) (a)	**1**	**6**	(b)	**1**	**6**
NISVAI	ŋa-cəkay	ru-cəkay	AULUA	boxol	drov-oxol
AXAMB	a-cəkay	rə-cəkay	NASVANG	i-cik	i-ləm-rə-cəkay
NĀTI	i-siʔ	seu-siʔ	PT. SANDWICH	cika	e-mo-cukay
AVETEIAN	i-sua	i-lav-sua	NAVWIEN	saxai	drax-saxal

The following, however, show significant differences between the form for 1 used on its own and the form for 1 used in the compound meaning 6. Again, I list them in two groups: Those in (10a) might just possibly be extreme examples of the case in (9b)—same origin but some phonological variation—where one could argue that only the first phoneme of the root meaning 1 seems to survive (and note *s > *-h* / *i,u _# is regular in Neve'ei). Those in (10b), however, show a quite different form for 1, uniformly *tes*. The form that apparently means 1 is bolded in each word for 6. (Avava *sapm* 1, *sou-t* 6 may belongs with (a) or (b).)

(10)(a)	**1**	**6**	(b)	**1**	**6**
TAPE	i-simək, i-sig	ləm-ji-**s**	UNUA	soxa	morov-**tes**
NASARIAN	i-sakawa	i-sau-**s**	BANAM BAY	soxa	maroc-**tes**
NEVE'EI	sevax	nsou-**h**	BWENELANG	boxol	marov-**tes**
LARËVAT	sag	nso-**s**	MASKELYNES	e-sua	(e)-məlev-**tes**
NAMAN	savax	nsou-**s**			
NEVERVER	i-sxam	i-jo-**s**			

The abbreviated form in (10a) is a feature of all Central-Western languages, plus their geographically closest neighbors, Tape and Nasarian. It is likely a Central-Western innovation that has later spread a little north and south. The replacement of the form for 1 with a different form, *tes*, in (10b) seems to have developed somewhere along the central east coast, and may suggest a closer relationship between Unua and Banam Bay with the Southeastern languages.

4.3 LIGATURE 6–9. I use the term "ligature" to refer to morphemes that are not themselves numerals but that occur in compound numerals: thus in Nasvang *i-ləm-**mə-rax**-ru* (3SG-5-**and-LIG**-2) 'seven', or in Neverver *naŋavul **nidruman** i-vas* (10 **LIG** 3SG-4) 'fourteen', the bolded items clearly link a numeral with another numeral. In many languages, the numerals 6–9 show a ligature but no preceding element for 5: thus Unua ***morov**-ru*, Nāti ***seu**-ru* '7' consist simply of 'LIG-2'.

The ligature linking tens and units (as *nidruman* in the Neverver example above) is a reflex of an earlier *dum^{w}a-na, and is found in just about every Malakula language; the only exceptions seem to be Maskelynes *pisan* (the verb meaning 'collect') and, apparently, *ay* in Navwien. It is thus of no interest to any internal subgrouping hypothesis.

As far as the ligature involved in compounds meaning 6–9 is concerned, in Lynch (2009) I described three ligatures that are reasonably widespread in the languages of Vanuatu. That discussion is summarized in table 1:

this gives the reconstructed form of the ligature, its probable (or possible) origin in Proto-North-Central Vanuatu (PNCV; see Clark 2009), as suggested in that 2009 paper, an example or two from Malakula, and its distribution in Vanuatu outside Malakula. Details of the distribution in Malakula are discussed below.

TABLE 1. COMMON LIGATURES IN VANUATU LANGUAGES AND THEIR ORIGINS

Protoform of ligature	Probable PNCV origin	Examples	Distribution
*lave-a	*lavi 'carry, take'	Aveteian *i-**lav**-rua* (3SG-LIG-2) = '7' Banam Bay ***rov**-ru* (LIG-2) = '7'	Torres, Banks, Maewo, Santo, Paamese, Ambrym, Shepherds, Efate
*[la]kau	*lakau 'cross over'	Nisvai ***rax**-rəl* (LIG-3) = '8' Avok ***axu**-rər* (LIG-3) = '8'	South Santo, Epi
*zau	[see below]	Nasarian *i-**sau**-vas* (3SG-LIG-4) = '9' Naman ***nso**-ves* (LIG-4) = '9'	(Malakula only)

Before examining the distribution of these ligatures in Malakula, I look briefly at the forms.

a. The form *lave-a as a variant of the verb *lavi 'fetch, get, carry, take' makes semantic sense: '5-get-2' for '7' has a certain logic. In addition, the forms of the ligature are identical with or very similar to the forms of the verb meaning 'take' in these languages, with often only a vocalic change; for example, Aveteian *lav* 'LIG', *lav* 'give'; Maskelynes *mə-lev* 'LIG', *lav-i* 'transfer'; Banam Bay *rov* 'LIG', *rav-i* 'take'.
b. For the origin of the second ligature in table 1, I had suggested the PNCV form *lakau 'cross over', which made semantic sense: '5-cross over-2' for '7'. However, while reflexes of this verb in all Malakula languages that have this ligature are *l*-initial (cf. Nisvai *ləkao*, Avok *lekao* 'go across, step over'), the ligatures themselves are *r*-, *dr*-, or zero-initial (see table 2), *never* *l*-initial, which suggests that they do *not* derive from *lakau. A possible alternative origin might be PNCV *raka-ti 'lift, raise, pull out', though I do not have any evidence to support or refute this. I refer to the form for the moment as *raka[ ].
c. In Lynch (2009), I suggested that *zau may derive from PNCV *sabo 'ignorant, incompetent, lost', some of whose reflexes mean 'other, different'—thus '5-other (hand)-2' for '7'. But I feel the formal and semantic differences are too great to support this speculation, and I have no alternative to suggest at this stage.

The distribution of the ligatures in Malakula is outlined in table 2.

TABLE 2. DISTRIBUTION OF 6–9 LIGATURES IN MALAKULA

	*lave-a	*raka[ ]	*zau	other
EASTERN	Unua *mo/rov-* Banam Bay *rov-*	Aulua *drox(u)-*		
Southeastern	Bwenelang *lov-* Maskelynes *mə/lev-*	Nasvang *mə-rax-* Nisvai *rax-* Port Sandwich *m-ox(u)-* Avok *axu-* Axamb *rax-*		
WESTERN Peripheral			Nāti *seu-*	Ninde *dumone-*[8]
Northwestern			V'ënen Taut *sa-* Tape *jə-*	
Southwestern	Lendamboi *lavu-* Aveteian *lav-*	Navwien *drax-* Naha'ai *ra-*	Avava *sou-* Nasarian *sau-* Nahavaq *sow-*	
Central-Western			Neve'ei *nsou-, nsu-* Larëvat *nso-* Naman *nso-, nsu-* Neverver *jo-*	

The form *lave-a has a transparent meaning, 'take, get, carry', and a wide and varied distribution both within Vanuatu and within Malakula. It gives the impression of being an "obvious" or natural development. That is, if you are going to replace monomorphemic '9', say, with something meaning '5 -X - 4', then a 'get' verb is a good candidate; and note especially here forms like Unua *mo-rov-*.or Maskelynes *mə-lev-*. in which the first element derives from *ma 'and': '5-and-get-4' for '9'. In addition, the verbs meaning 'get' in these languages are very similarly phonologically to the ligature. All of this, however, suggests that this may have been an

[8] In Ninde, the same ligature is used for 6–9 and for linking tens and units: *dumone-ves* (LIG-4) '9', *nalaŋal-dumone-ves* (10-LIG-4) '14'.

independent development that took place a number of different times in Vanuatu, and offers very little to subgrouping hypotheses within Malakula.

The ligature *raka[ ] may not be connected at all to *lakau 'cross over', and may thus not be related to similar forms in South Santo and Epi. It is largely confined to the Southeastern languages (plus Aulua), from where it seems to have spread across the linkage boundary to Navwien and thence to Naha'ai. It thus offers tentative support for the Southeastern grouping.

The ligature *zau is perhaps of most use here. A glance at its distribution would suggest that it was probably an early Western linkage innovation that spread into all the component subklinkages, but which was replaced —either through contact or by later innovation—by *lavea in Aveteian and Lendamboi, *raka[ ] in Navwien and Naha'ai, and by the form *dum^{w}a-na that links tens and units in Ninde. I make this suggestion because, if it was a fairly recent development, one might expect to find some obvious semantic origin (as with *lavea, for example); but there is no such obvious origin, which suggests it is a form of some antiquity.

4.4 SUMMARY. The distribution of the innovations discussed in this section are summarized in table 3. A double line separates the three major groupings, while single lines separate subgroups or sublineages within these major groupings. Unless otherwise noted, within that table, ✓ means 'participates in the innovation', a blank means 'does not participate in the innovation', and – means 'no data available'. At the top of each column is a code labeling the innovations, as follows:

6–9: the numerals 6–9 are compounds of 1-4; 10 derives from *saŋapulu(q): see 4.1
20: the form for 20 is based on 'man, person', not on a compound of 2 and 10: see (4)
N2–4: 2–4 show the nasalized reflex of the initial C but 7–9 show the regular oral reflex: see (5)
4–N9: forms for 9 show the nasalized reflex of the initial C but not forms for 4: see (6)
49C# final C in 4 differs from that in 9; † = apparent contrast between *pati (in 4) and *pat (in 9): see (7)
10*l- *saŋapulu(q) '10' irregularly > *laŋapulu(q): see 4.2.2, (8)
6≠1 the form for 1 in the compound 6 is different from 1, either partially (a) or totally (b)
LIG reflexes of the three ligatures given in table 2 are listed and color-coded. Here, x means some other form is used and blank means no ligature (true decimal)

TABLE 3. DISTRIBUTION OF NUMERAL INNOVATIONS

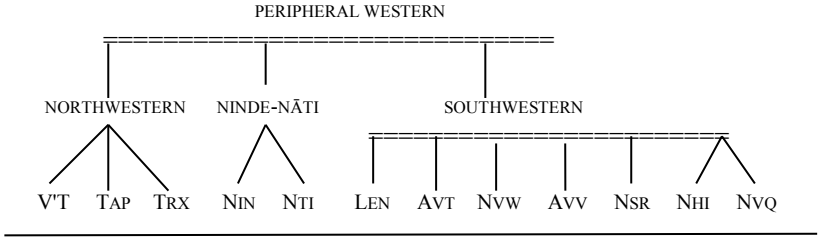
				6–9	20	N2–4	4–N9	49C#	10*l-	6≠1	LIG
			Malua Bay								
N			Nese								
T	N Coast		Botovro		–						
H			Vovo		–						
			Vao		–						
			Uripiv								
			Aulua	✓	–	✓					raka[ ]
E			Unua	✓			✓			✓b	lavea
A			Banam Bay	✓			✓	✓		✓b	lavea
S			Bwenelang	✓		✓		✓		✓b	lavea
T			Nasvang	✓	✓		✓				raka[ ]
E	South-		Nisvai	✓			✓				raka[ ]
R	eastern		Port Sandwich	✓			✓				raka[ ]
N			Avok	✓							raka[ ]
			Axamb	✓	✓						raka[ ]
			Maskelynes	✓			✓			✓b	lavea
			Ninde	✓					✓		x
			Nāti	✓	✓			✓	✓		zau
		North-	V'ënen Taut	✓				†			zau
		western	Tape	✓				†		✓a	zau
W	Peripheral		Tirax								
E			Aveteian	✓	–				✓		lavea
S		South-	Lendamboi	✓	✓				✓		lavea
T		western	Navwien	✓	✓	✓			✓		raka[ ]
E			Naha'ai	✓	✓				✓		raka[ ]
R			Nahavaq	✓	✓			✓	✓		zau
N			Avava	✓					✓		zau
			Nasarian	✓					✓	✓a	zau
	Central-		Neve'ei	✓						✓a	zau
	Western		Larëvat	✓	–					✓a	zau
			Naman	✓						✓a	zau
			Neverver	✓						✓a	zau

5. IMPLICATIONS. It will be immediately clear from table 3 that none of these innovations have occurred within the Northern Malakula subgroup, since all of them (except for *saŋapulu(q) > *laŋapulu(q)) could only arise in an imperfect decimal or a mixed system. We can thus ignore those languages (except in any cases where they may have influenced languages of other groupings).

As far as other groupings are concerned, the following points can be made:

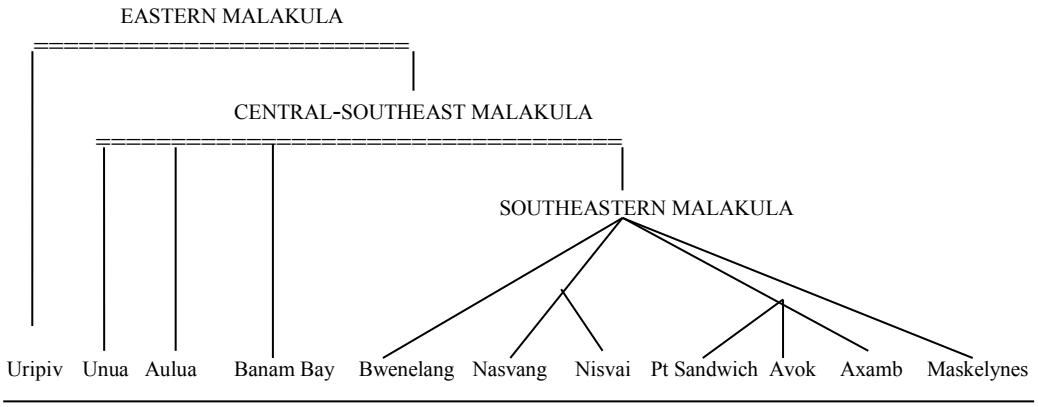
- If my hypothesis in 4.3 regarding the ligature *zau is correct, this would provide confirmation of the validity of the Western Malakula linkage, which was only fairly weakly established.
- The use of an abbreviated form for '1' in the numeral '6' (4.2.3) adds some strength to the previously weak Central-Western sublinkage.
- The development of the vigesimal unit '20' from forms for 'man, person' and the irregular change *s > *l in *saŋapulu(q) '10' both provide solid confirmation of the validity of the Southwestern sublinkage. Nāti shows both innovations, and Ninde shows the second. In terms of strict phonological innovations, Ninde and Nāti appear closer to their Northwestern relatives, especially V'ënen Taut and Tape; in terms of lexical irregularities, they seem to be closer to their Southwestern relatives. Ignoring for the moment the uncertain position of Tirax, I am inclined to slightly revise the subgrouping hypothesis for the Peripheral Western sublinkage as figure 4.

FIGURE 4. PERIPHERAL WESTERN SUBLINKAGE (REVISED)



- A number of innovations are found within the Southeastern sublinkage, which provides confirmation for that hypothesis. In two cases—nasalized reflex of the form for '4' in '9' (4.2.1) and the use of a totally different form for 1, *tes*, in '6' (4.2.3)—the innovations are shared with Unua and Banam Bay; while the ligature *raka[ ] is shared with Aulua. These three languages are geographically to the north of the Southeastern linkage, and remain isolated from the remaining eastern language, Uripiv, by a number of intrusive Central-Western languages. There may be some evidence here to modify the internal structure of the Eastern Malakula linkage as in figure 5.

FIGURE 5. EASTERN MALAKULA LINKAGE (REVISED)



There is some evidence for features dispersing through contact, though none of it is very strong. It is possible (a) that Tirax and Uripiv borrowed the pure decimal system from a northern language, likely Vao (4.1.1); (b) that Nisvai and Axamb borrow innovation of using forms for 'man, person' to mean '20' from a Southwestern language (4.1.2); (c) that Tape and Nasarian the construction for '6' with a phonologically abbreviated form for '1' from a Central-Western language (4.2.3); and (d) that Navwien and Naha'ai borrowed the ligature *raka[ ] from a Southeastern language (4.3).

6. CONCLUSION. In this paper, I have described—in what I believe is a more complete and systematic way than Charpentier (1987)—the historical development of the numeral systems of the modern Malakula languages. Although some features spread through contact, there is quote a bit of evidence showing that certain features developed early and were inherited in a group of daughter-languages. The data outlined here confirm the general outline of the internal subgrouping hypothesis given in Lynch (to appear), and also add some refinements to that hypothesis, especially in relation to the position of Ninde and Nāti.

APPENDIX. DATA SOURCES

Language	Data sources	Language	Data sources
Malua Bay	Tryon (1976)	Nasvang	Charpentier (1982, 1987)
Nese	Crowley (2006a)	Maskelynes	Healey (2013)
Vovo	Tryon (1976)	Avok	Charpentier (1982, 1987)
Botovro	Tryon (1976)	Axamb	Charpentier (1982, 1987)
Vao	Tryon (1976)	Lendamboi	Charpentier (1982, 1987)
Uripiv	McKerras (2001)	Naha'ai	Charpentier (1982, 1987)
Tirax	Amanda Brotchie	Nāti	Crowley (1998)
Naman	Crowley (2006b)	Nahavaq	Dimock (2009)
Neverver	Barbour (2012)	Navwien	Charpentier (1982, 1987)
Unua	Pearce (2015)	Ninde	Charpentier (1982, 1987), Dimock et al. (n.d.)
Aulua	Charpentier (1982, 1987)	Aveteian	Charpentier (1982, 1987)
Nasarian	Charpentier (1982, 1987)	Neve'ei	Jill Musgrave
Banam Bay	Charpentier (1982, 1987)	Avava	Crowley (2006c)
Bwenelang	Charpentier (1982, 1987)	Larëvat	Terry Crowley
Nisvai	Charpentier (1982, 1987)	Tape	Crowley (20006d)
Port Sandwich	Charpentier (1982, 1987)	V'ënen Taut	Fox (1979)

REFERENCES

Barbour, Julie. 2012. *A grammar of Neverver*. Berlin/Boston: Walter de Gruyter.

Charpentier, Jean-Michel. 1982. *Atlas linguistique du sud Malakula / Linguistic atlas of south Malakula*. Paris: SELAF.

———. 1987. La numeration au Sud-Malakula (Vanuatu). In *A world of language: Papers presented to Professor S. A. Wurm on his 65th birthday*, ed. by Donald C. Laycock and Werner Winter, 105–18. Canberra: Pacific Linguistics.

Clark, Ross. 1985. ———. 2009. **Leo tuai: A comparative lexical study of North and Central Vanuatu languages*. Canberra: Pacific Linguistics.

Crowley, Terry. 1998. A salvage sketch of Nāti (southwest Malakula, Vanuatu). In *Papers in Austronesian Linguistics No. 5*, ed. by Darrell Tryon, 101–48. Canberra: Pacific Linguistics.

———. 2006a. *Nese: A diminishing speech variety of Northwest 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. *The Avava language of Central Malakula (Vanuatu).* [Ed. by John Lynch.] Canberra: Pacific Linguistics.

———. 2006d. *Tape: A declining language of Malakula (Vanuatu).* [Ed. by John Lynch.] Canberra: Pacific Linguistics.

Dimock, Laura Gail. 2009. A grammar of Nahavaq (Malakula, Vanuatu). PhD thesis, Victoria University of Wellington.

Dimock, Laura, Emily Greenbank, Leina Isno, and Elizabeth Pearce. n.d. A sketch grammar of Ninde. MS, Victoria University of Wellington.

Fox, G. J. 1979. *Big Nambas grammar*. Canberra: Pacific Linguistics.

Healey, David Stephen. 2013. A grammar of Maskelynes: the language of Uluveu Island, Vanuatu. PhD thesis, University of the South Pacific.

Huffman, Kirk W. 1996. Trading, cultural exchange and copyright: Important aspects of Vanuatu arts. In *Arts of Vanuatu*, ed. by Joël Bonnemaison, Christian Kaufmann, Kirk Huffman, and Darrell Tryon, 182–94. Bathurst: Crawford House Publishing.

Lynch, John. 2008. Liquid vocalization and loss in Central Vanuatu. *Oceanic Linguistics* 47:294-315.

———. 2009. At sixes and sevens: The development of numeral systems in Vanuatu and New Caledonia. In *Discovering history through language: Papers in honour of Malcolm Ross*, ed. by Bethwyn Evans, 391–411. Canberra: Pacific Linguistics.

———. To appear. Malakula internal subgrouping: Phonological evidence. *Oceanic Linguistics* 55(2), 2016.

Lynch, John, and Amanda Brotchie. 2010. Vowel loss in Tirax and the history of the apicolabial shift. *Oceanic Linguistics* 49:369–88.

Lynch, John, and Terry Crowley. 2001. *Languages of Vanuatu: A new survey and bibliography*. Canberra: Pacific Linguistics.

Lynch, John, Malcolm Ross, and Terry Crowley. 2002. *The Oceanic languages*. Richmond, Surrey: Curzon Press.

McKerras, Ross. 2001. Uripiv dictionary. MS.

Pearce, Elizabeth. 2015. *A grammar of Unua*. Boston/Berlin/Munich: Walter de Gruyter.

Tryon, D. T. 1976. *New Hebrides languages: An internal classification*. Canberra: Pacific Linguistics.